

Mechanisms of
eukaryotic gene expression
on a single molecule level:
From transcription initiation
to nucleosome remodeling

Barbara Treutlein



DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER FAKULTÄT CHEMIE DER
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Mechanisms of eukaryotic gene expression on a single molecule level:

From transcription initiation to nucleosome remodeling

A single molecule FRET and Nano-Positioning System study



Barbara Treutlein

aus

Reutlingen, Deutschland

2012

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-8439-0591-6

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Jens Michaelis betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 05. Juli 2012

(Barbara Treutlein)

Dissertation eingereicht am: 05. Juli 2012
1. Gutachter: Prof. Dr. Jens Michaelis
2. Gutachter: Prof. Dr. Patrick Cramer
Mündliche Prüfung am: 25. Juli 2012

© Verlag Dr. Hut, München 2012
Sternstr. 18, 80538 München
Tel.: 089/66060798

Die Informationen in diesem Buch wurden mit großer Sorgfalt erarbeitet. Dennoch können Fehler, z.B. bei der Beschreibung des Gefahrenpotentials von Versuchen, nicht vollständig ausgeschlossen werden. Verlag, Autoren und ggf. Übersetzer übernehmen keine juristische Verantwortung oder irgendeine Haftung für eventuell verbliebene fehlerhafte Angaben und deren Folgen.

Alle Rechte, auch die des auszugsweisen Nachdrucks, der Vervielfältigung und Verbreitung in besonderen Verfahren wie fotomechanischer Nachdruck, Fotokopie, Mikrokopie, elektronische Datenaufzeichnung einschließlich Speicherung und Übertragung auf weitere Datenträger sowie Übersetzung in andere Sprachen, behält sich der Autor vor.

1. Auflage 2012

Meinen Eltern

Abstract

Eukaryotic gene expression begins with transcription of genomic DNA into messenger RNA, carried out by the multi-subunit enzyme RNA polymerase II (Pol II) with the help of numerous transcription factors. Transcription initiation requires recruitment of Pol II to the promoter DNA by bound transcription factors, however accessibility to genomic DNA is hindered due to its compaction into nucleosomes. Therefore, cells have evolved a set of ATP-hydrolyzing enzymatic complexes, chromatin remodelers, that reposition nucleosomes, thereby providing regulated access to the underlying DNA.

This thesis focuses on understanding the molecular mechanisms of transcription initiation and nucleosome remodeling. Since both processes are characterized by large conformational changes and a high flexibility of nucleoprotein complexes, standard high-resolution structural methods are hindered and instead direct visualization in real time is required as provided by single molecule techniques.

In the first part of this work, I used single molecule fluorescence resonance energy transfer (smFRET) experiments, Nano-Positioning System (NPS) analysis and x-ray crystallographic information to determine the three-dimensional architecture of a minimal Pol II open promoter complex (OC) consisting of promoter DNA, TBP, Pol II and general transcription factors TFIIB and TFIIF. In the OC, TATA-DNA and TBP reside above the Pol II cleft between clamp and protrusion domains. The TFIIB core domain is displaced from the Pol II wall, where it is located in the closed promoter complex. Furthermore, I directly observed the downstream DNA to be dynamically loaded into and unloaded from the Pol II cleft at a timescale of seconds. These results uncover large overall structural changes during the initiation-elongation transition.

In the second part, I applied the same experimental approach to determine the location of the three constituting domains of chromatin remodeler Chd1 (Chromodomain-helicase-DNA-binding protein 1) in a Chd1-nucleosome complex that exhibited intrinsic dynamics. The NPS results allowed me to construct a preliminary model of the Chd1-nucleosome complex, in which the DNA-binding domain is associated with extranucleosomal DNA at the nucleosome entry site, the tandem chromodomains are located below the entry site close to histone H4 tail and the ATPase motor binds nucleosomal DNA between dyad and superhelical location +1. Furthermore, I used smFRET to follow in real-time the structural dynamics of nucleosomal DNA during Chd1 catalyzed repositioning. FRET time trajectories revealed gradual and bidirectional translocation of nucleosomal DNA by Chd1 and the data allowed me to propose a model for the remodeling mechanism of Chd1, which involves formation and propagation of a DNA loop.

Zusammenfassung

Eukaryotische Genexpression beginnt mit der Transkription von genomischer DNS zu Boten-RNS durch die RNS Polymerase II (Pol II), welche die Hilfe zahlreicher Transkriptionsfaktoren benötigt. Initiation der Transkription erfordert die Rekrutierung der Pol II zum Gen-Promoter durch dort gebundene Transkriptionsfaktoren. Jedoch ist der Zugang zur genomischen DNS erschwert durch ihre Verpackung in Nukleosomen. Um regulierten Zugang sicher zu stellen sind Zellen mit ATP-hydrolysierenden enzymatischen Komplexen, so genannten *Chromatin Remodelers* ausgestattet, die in der Lage sind Nukleosomen entlang der DNS zu verschieben.

Das Ziel dieser Dissertation ist es die molekularen Mechanismen der Transkriptions-Initiation und der Nukleosom-Repositionierung aufzuklären. Beide Prozesse sind durch große Konformationsänderungen und eine hohe Flexibilität der involvierten Protein-DNS Komplexe charakterisiert, so dass Strukturaufklärung mittels hochauflösender Standard-Methoden (z.B. Röntgen-Kristallographie) erschwert ist. Anstelle dessen ist eine direkte Visualisierung in Echtzeit erforderlich, so wie es Einzelmolekül-Methoden bieten.

Im ersten Teil dieser Arbeit wurde mit Hilfe von Fluoreszenz-Resonanz-Energietransfer (FRET) Experimenten einzelner Moleküle, *Nano-Positioning System* (NPS) Analyse und röntgen-kristallographischen Informationen die dreidimensionale Architektur des offenen Pol II Promoterkomplexes (*Open Complex*, OC) bestehend aus Pol II, Promoter-DNS, TBP und Transkriptionsfaktoren TFIIB und TFIIF aufgeklärt. Dabei konnte der dynamische Ladungsprozess der Transkriptionsblase in die Pol II *cleft* auf einer Zeitskala von Sekunden direkt beobachtet werden. Die Ergebnisse klären umfassende Konformationsumwandlungen auf, die während des Übergangs von der Initiations- zur Elongationsphase der Transkription stattfinden.

Im zweiten Teil wurde dieselbe Methodik verwendet, um die unterschiedlichen Domänen des Remodelers Chd1 (Chromodomain-helicase-DNA-binding protein 1) relativ zu einem Nukleosom zu positionieren. Die untersuchten Chd1-Nukleosom Komplexe zeigten strukturelle Flexibilität. Mit Hilfe der NPS Ergebnisse konnte ein vorläufiges Modell für die Architektur eines Chd1-Nukleosom Komplexes erstellt werden. Darüber hinaus wurden sm-FRET Experimente verwendet, um die strukturellen Veränderungen der nukleosomalen DNS während Repositionierung durch Chd1 in Echtzeit zu verfolgen. Die beobachteten dynamischen FRET Trajektorien wurden mit Hilfe von *Hidden Markov Modeling* Analyse quantitativ untersucht. Auf Basis der Ergebnisse konnte ein Modell für den Mechanismus der Repositionierung erstellt werden, der die Ausbildung und das Propagieren einer DNS Schleife involviert.

Contents

Contents	xi
Introduction	1
1. Transcription and Chromatin	7
1.1. Structure and function of RNA polymerase II	7
1.1.1. The RNA polymerase II transcription cycle	7
1.1.2. Structure of the Pol II elongation complex	9
1.1.3. Specific function of initiation factors	11
1.1.4. Single molecule studies of transcription	11
1.2. Chromatin structure and dynamics	15
1.2.1. Chromatin organization in eukaryotes	15
1.2.2. Nucleosome stability and dynamics	17
1.2.3. Chromatin remodeling complexes	18
1.2.4. The CHD family of chromatin remodelers	19
1.2.5. Chromodomain-helicase-DNA-binding protein 1 - Chd1	20
1.2.6. Mechanisms of nucleosome repositioning and spacing	23
1.2.7. Single molecule studies of nucleosome remodeling	26
2. Physical basics	29
2.1. Fluorescence	29
2.1.1. Basics	29
2.1.2. Fluorescence anisotropy	32
2.1.3. Fluorescence Resonance Energy Transfer (FRET)	32
2.1.4. Förster distance	34
2.1.5. Site-specific fluorescent labeling of proteins	36
2.1.6. Single molecule fluorescence	37
2.2. Total internal reflection fluorescence microscopy (TIRFM)	38
2.2.1. Experimental setup	40
2.2.2. Preparation of flow chamber for smFRET measurements	42
2.3. Acquisition, processing and analysis of smFRET data	44
2.4. Hidden Markov Modeling (HMM)	48

Contents

2.5. Bayesian data analysis and the Nano-Positioning System	51
2.5.1. Probability calculus	52
2.5.2. Bayesian parameter estimation	53
2.5.3. Bayesian parameter estimation in the global NPS	53
3. Dynamic architecture of a minimal Pol II open promoter complex	59
3.1. Abstract	59
3.2. Introduction	60
3.3. Results	61
3.3.1. Assembly of OCs with labeled DNA and factors	61
3.3.2. Downstream DNA occupies the Pol II cleft	64
3.3.3. Dynamic loading of downstream DNA into the cleft	64
3.3.4. TFIIF stabilizes downstream DNA in Pol II cleft	67
3.3.5. TBP and TATA DNA reside above the cleft	70
3.3.6. Upstream DNA is stabilized by initiation factors	70
3.3.7. Model of the OC	71
3.3.8. B-core is displaced from the wall	72
3.4. Discussion	76
3.5. Summary and Outlook	79
3.6. Experimental Procedures	82
3.6.1. Protein expression, purification and labeling	82
3.6.2. Nucleic Acid scaffolds	83
3.6.3. Open and elongation complex assembly	83
3.6.4. Electrophoretic mobility shift assay	84
3.6.5. <i>In vitro</i> transcription assay	84
3.6.6. smFRET measurements: Experimental setup and data analysis . .	84
3.6.7. Hidden Markov Model (HMM) analysis of dynamic smFRET time traces	85
3.6.8. Global Nano-positioning system analysis	85
3.6.9. Modeling	88
4. Single molecule FRET studies of nucleosome remodeling by Chd1	89
4.1. Abstract	89
4.2. Introduction	90
4.3. Probing Chd1 induced nucleosome dynamics by smFRET	91
4.3.1. Results	91
4.3.2. Discussion	112
4.4. Architecture of a Chd1-nucleosome complex	119
4.4.1. Results	119
4.4.2. Discussion	135
4.5. Summary and Outlook	139

4.6. Experimental Procedures	142
4.6.1. Buffers	142
4.6.2. Nucleosomal DNA	143
4.6.3. LCR preparation of labeled nucleosomal DNA	144
4.6.4. PCR preparation of labeled nucleosomal DNA	145
4.6.5. SEC purification of labeled nucleosomal DNA	146
4.6.6. Unlabeled nucleosomal and competitor DNA	147
4.6.7. Preparation of Histone Octamers	148
4.6.8. Nucleosome assembly via salt gradient dialysis	148
4.6.9. Expression and Purification of Chd1	150
4.6.10. Labeling of Chd1	152
4.6.11. Chd1 Binding assay	152
4.6.12. Nucleosome remodeling assay	153
4.6.13. Oxygen-scavenging system and Trolox	154
4.6.14. smFRET analysis of nucleosome dynamics induced by Chd1	155
4.6.15. Hidden Markov Modeling (HMM) analysis of nucleosome dynamics	156
4.6.16. Chd1 positioning experiments	157
4.6.17. Global NPS analysis of a Chd1-nucleosome complex	157
5. Outlook	163
A. Supplemental Data	169
A.1. NPS analysis of Pol II OC	169
A.2. HMM analysis of dynamic FRET traces during Chd1 remodeling	175
A.3. NPS analysis of Chd1-nucleosome complex	180
Bibliography	I

Introduction

The polymeric deoxyribonucleic acid (DNA) of every cell encodes information that is essential to the development and maintenance of life. This information is passed on during gene expression from the DNA via ribonucleic acid (RNA) to proteins, as first enunciated in the central dogma of molecular biology by Francis Crick in 1958 (published in Nature in 1970 [1]). The underlying fundamental biological processes, namely DNA replication, transcription and translation, are found in all three kingdoms of life (Figure 0.0.1).

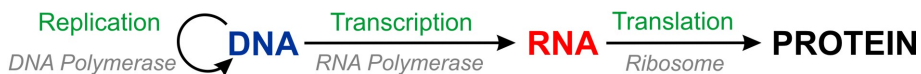


Figure 0.0.1.: Central dogma of molecular biology describing the directional flow of information from DNA via RNA to protein [1].

Transcription of protein-coding genes into messenger RNA (mRNA) is a complex and intricately regulated process that controls cell growth and differentiation [2]. It proceeds through multiple steps and is customarily divided into three stages designated initiation, elongation and termination.

In eukaryotes, the synthesis of mRNA is carried out by RNA Polymerase II (Pol II) along with a large set of protein factors that specifically interact with Pol II during the different stages and oftentimes possess regulatory functions. Most importantly, Pol II requires the assistance of general transcription factors (GTFs) TFIIB, TFIID, TFIIIE, TFIIF and TFIIH, which localize the enzyme at proper promoter sequences and enable transcription initiation by facilitating conformational changes essential to Pol II activity. During transcription initiation, Pol II is recruited to the promoter by GTFs followed by formation of the closed promoter complex (CC) [3, 2, 4]. In addition to GTFs, recruitment of Pol II to the promoter is greatly influenced by the Mediator complex, and DNA-binding transcription activators. Subsequently, DNA surrounding the transcription start site (TSS) is melted by the helicase activity of TFIIH and inserted into the active center cleft of Pol II. The resulting open promoter complex (OC) enables RNA synthesis, which triggers release of the GTFs and conversion of the OC to a stable elongation complex (EC).

Transcription requires genomic DNA to be accessible to the transcription machinery to bind the promoter and translocate along the gene during mRNA synthesis. This is provided in bacteria, where DNA exists in its bare form. However in eukaryotes, genomic

DNA is packaged into chromatin through the interaction with histone proteins, which occludes stretches of DNA thereby creating obstacles to all DNA associated processes such as transcription, replication and DNA repair. The basic unit of chromatin is the nucleosome, formed by wrapping of about 147 bp of DNA in 1.65 turns around an octameric complex of core histones, a central (H3-H4)₂ hetero-tetramer and two peripheral H2A-H2B dimers [5]. Arrays of nucleosomes are further compacted into higher order structures such as the 30 nm chromatin fiber. On one hand, chromatin restricts inappropriate access to DNA and allows condensation and organization of the ~2 m long genomic DNA such that it fits the microscopic dimensions of a cell nucleus, ~10 μ m in diameter. However, at the same time chromatin provides another layer of regulation since it can actively participate in controlling access to the underlying genomic information. To this aim, cells are equipped with a set of specialized chromatin remodeling complexes that use the energy of ATP hydrolysis to modify the structure of chromatin. Remodelers reposition, destabilize, eject or restructure nucleosomes, thereby allowing rapid and regulated access to the underlying genomic DNA [6]. Remodelers are recruited to specific loci on chromatin where their action is needed through covalent modifications of histone tails, so-called histone marks. These marks are placed by histone modifying enzymes, which in turn are recruited by Pol II itself via its long C-terminal domain (CTD) and CTD-associated factors [7]. This manifests a strong interplay between chromatin remodeling and transcription and shows that an understanding of the mechanisms of chromatin remodeling is crucial for a better understanding of transcriptional regulation.

My work focuses on the two processes of transcription initiation and nucleosome repositioning. Both processes have been extensively studied using functional biochemical assays as well as structural analysis mainly by x-ray crystallography or electron microscopy, which has greatly advanced our understanding of the underlying mechanisms. However, many open questions remain and the applied traditional methods are limited in finding answers, because of the size, complexity, heterogeneity and flexibility of transcription initiation intermediates and remodeler-nucleosome complexes. Biochemical analysis has the limitation of monitoring the properties of large ensembles, thereby obscuring individual behaviors and occluding heterogeneity that might exist between different populations of molecules. X-ray crystallography requires the complex of interest to be captured in a crystal, which is impossible if complexes within a sample are heterogeneous in composition or conformation or if they contain too many flexible domains. Finally, traditional structural methods only provide a static picture of otherwise dynamic processes and hence cannot be used to study structural dynamics.

Both processes, initiation of transcription and nucleosome remodeling involve large multi-protein complexes and are characterized by large conformational changes. Therefore, analysis of underlying mechanisms requires the direct visualization of individual complexes in real time as provided by single molecule techniques.

Single molecule Förster resonance energy transfer (smFRET) is ideally suited to struc-

turally analyze flexible domains and conformational changes within macromolecular complexes, because it allows to measure distances in real-time between two dye molecules attached to the complex [8, 9]. Amongst others, it has been successfully applied to the analysis of transcription initiation complexes of the bacterial [10, 11] and mitochondrial [12] RNA polymerase and to the study of nucleosome dynamics [13] and remodeling [14]. By trilateration of several smFRET derived distances, a region of unknown position can be located with respect to known positions within a complex [15, 16, 17, 18, 19]. However, experimental uncertainties associated with smFRET measurements affect the determined most likely positions and must be accounted for by computing a three-dimensional probability density function (PDF) for each position. The recently established Nano-Positioning System (NPS) computes such PDFs using smFRET data in combination with x-ray crystallographic information and Bayesian parameter estimation [20].

NPS analysis of Pol II transcription elongation complexes has revealed the position of the 5'-end of exiting RNA [15], the influence of TFIIB on RNA position [20], and the course of nontemplate and upstream DNA [21]. NPS was recently extended to include a global data analysis to improve localization accuracy (global NPS), and the docking of macromolecules of known structure [22].

This thesis

In this thesis, I apply smFRET experiments and global NPS analysis to explore the mechanism of two fundamental processes during eukaryotic gene expression: transcription initiation by *S.cerevisiae* RNA Pol II and nucleosome repositioning by the *S.cerevisiae* remodeler Chd1 (Chromodomain-helicase-DNA-binding protein 1).

In the first part, I used smFRET and global docking NPS analysis to uncover the dynamic architecture of a minimal Pol II OC consisting of promoter DNA, TBP, TFIIB, TFIIF and Pol II. The results uncovered large overall structural changes during the transcription initiation-elongation transition and an intrinsically dynamic nature of the Pol II OC.

In the second part, I applied smFRET to explore the structural dynamics of nucleosomes during binding and remodeling by individual Chd1 remodelers in real-time. Processive, bidirectional translocation of nucleosomal DNA was observed and a remodeling mechanism was proposed that involves formation and propagation of a DNA loop over the histone octamer surface. Moreover, using smFRET together with global docking NPS analysis, I examined the architecture of a Chd1-nucleosome complex. Intrinsic dynamics were observed and the different functional domains of Chd1 could be mapped to different parts of the nucleosome.

In summary, the studies presented in this thesis shed light on the molecular mechanisms underlying transcription initiation and nucleosome remodeling and present a significant contribution to a better understanding of both processes. Moreover, they emphasize the suitability and great potential of smFRET and global NPS analysis to the structural and mechanistic investigation of complex and dynamic processes of gene expression.

Organization of the chapters

- In chapter 1, I introduce the biological background of this thesis by presenting the basics of transcription by eukaryotic RNA polymerase II, chromatin structure and dynamics as well as chromatin remodeling. Besides, I briefly present single molecule studies of transcription and nucleosome remodeling.
- In chapter 2, I introduce the physical background of the methods and phenomena applied in this thesis: Fluorescence, Förster resonance energy transfer, total internal reflection fluorescence microscopy (TIRF), Hidden Markov Modeling (HMM) and Bayesian data analysis.
- In chapter 3, I present my work on transcription initiation. In this study smFRET experiments and global NPS analysis are applied to uncover the dynamic architecture of a minimal Pol II open promoter complex.
- In chapter 4, I focus on nucleosome repositioning catalyzed by the chromatin remodeling complex Chd1. I present smFRET experiments that elucidate the structural dynamics of nucleosomes induced by Chd1 as well as smFRET experiments combined with global NPS analysis that decipher the architecture of a Chd1-nucleosome complex.
- An outlook (chapter 5) on future experiments concludes this thesis.

Publications of thesis work

- *Dynamic architecture of a minimal RNA Polymerase II open promoter complex.*
B. Treutlein, A. Muschielok, J. Andrecka, A. Jawhari, C. Buchen, D. Kostrewa, F. Hoeg, P. Cramer and J. Michaelis,
Molecular Cell **46**, 136-146 (2012).
- *Direct observation of single RNA polymerase processing through a single endogenous gene in a living yeast cell.*
B. Treutlein and J. Michaelis,
Angewandte Chemie International Edition **50**, 9788-9790 (2011).
- *Nano-Positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex.*
J. Andrecka, B. Treutlein, M. A. I. Arcusa, A. Muschielok, R. Lewis, A. C. M. Cheung, P. Cramer and J. Michaelis,
Nucleic Acids Research **37**, 5803-5809 (2009).

1. Transcription and Chromatin

1.1. Structure and function of RNA polymerase II

Transcription is the first and most highly regulated step in gene expression. It is carried out by DNA-dependent RNA polymerases that use genomic DNA as template for the synthesis of RNA. The existence of RNA polymerases was first discovered independently by Jerard Hurwitz [23] and Samuel Weiss [24] in 1960. While prokaryotes have only one RNA polymerase responsible for transcription of all genes, transcription in eukaryotes is carried out by three different RNA polymerases: Pol I, Pol II and Pol III. Moreover, a fourth type of RNA polymerase, Pol IV, exists in plants [25, 26]. Each polymerase is responsible for the transcription of a specific set of RNAs. Pol I synthesizes 5.8S, 18S and 28S ribosomal RNAs, Pol II produces messenger RNAs from all protein-coding genes and small nuclear RNAs and Pol III synthesizes transfer RNAs and other small RNAs. All three polymerases are multisubunit protein complexes and comprise 14 (Pol I, 589 kDa), 12 (Pol II, 512 kD) and 17 (Pol III, 693 kD) subunits [27]. 10 subunits form the enzyme core, with which the remaining peripheral subunits interact.

1.1.1. The RNA polymerase II transcription cycle

Transcription of protein-coding genes into messenger RNA (mRNA) is a complex and intricately regulated process that controls cell growth and differentiation [2]. It proceeds through a number of distinct, well ordered steps that form the so-called *transcription cycle* (Figure 1.1.1) [28, 2]. The transcription cycle can be roughly divided into three stages designated initiation, elongation and termination.

Transcription initiation

Transcription initiation at eukaryotic protein-coding gene promoters is a multi-step process that is highly regulated through the involvement of over a hundred of protein factors [29, 30]. First, gene specific activators recognize and bind to enhancer motives that are mainly located upstream of the gene within an upstream-activating sequence. Oftentimes, a gene is controlled by the combined action of multiple activators, which increases the level of regulation. Activators then recruit coactivators such as the large 20-subunit SAGA complex or the 25-subunit Mediator as well as chromatin remodelers and histone modify-

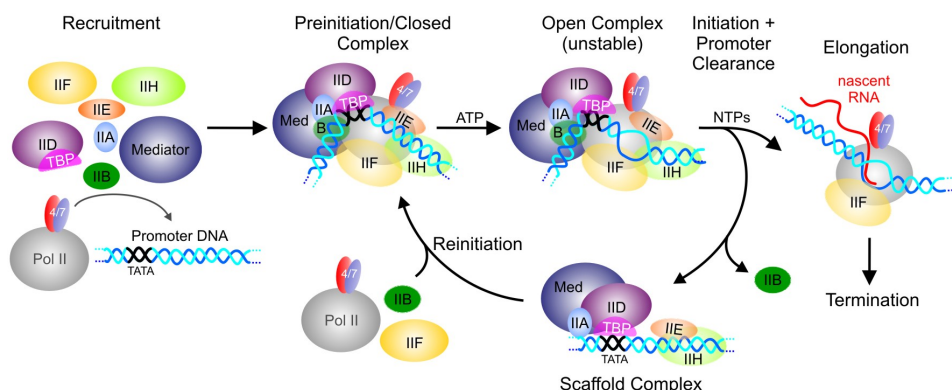


Figure 1.1.1.: The RNA polymerase II transcription cycle.

ing enzymes, such as ISWI, RSC, SWI/SNF, RSC, INO80, CHD (see section 1.2.3). Those enzymes alter the structure of chromatin to make the core promoter accessible [7, 31]. Subsequently, general transcription factors (GTFs) TFIIA, -B, -D, -E, -F and -H assemble at the promoter and recruit Pol II to form the pre-initiation or closed promoter complex (PIC/CC) [3, 2, 4]. Pol II, TFIIB, TFIIF, and the TFIID subunit TATA box-binding protein (TBP) suffice to form a minimal CC [32]. Upon CC formation, promoter DNA surrounding the transcription start site (TSS) is melted by the ATP-hydrolyzing helicase domains of TFIIE and inserted into the active center cleft of the polymerase resulting in the open promoter complex (OC). The minimal OC sufficient for promoter-dependent transcription *in vitro* is constituted of Pol II, TBP and TFIIB and is stimulated by the presence of TFIIF [33] (see section 1.1.3 for details about the function of TFIIB, TBP and TFIIF). The OC enables RNA synthesis, which triggers release of the general TFs and conversion of the OC to a stable elongation complex (EC). With the help of TFIIB, Pol II finds the TSS and starts mRNA synthesis. When the nascent transcript exceeds a length of about seven nucleotides, Pol II enters the elongation phase and escapes the promoter (*promoter clearance*) [34]. Most of the initiation factors are left behind as so-called promoter-bound scaffold, which greatly facilitates reinitiation by another Pol II [2].

Transcription elongation

During the elongation phase, Pol II moves along the gene with single nucleotide steps and processively produces an RNA transcript, complementary to template DNA. This precursor mRNA product is further processed by capping and splicing enzymes in a co-transcriptional manner (see below) to yield the final mRNA template for protein synthesis. Promoter-proximal pausing and backtracking of Pol II can interrupt the elongation process and elongation factors such as TFIIS are required to reactivate Pol II [35]. The long C-terminal domain (CTD) of Pol II Rpb1, which consists of multiple heptapeptide (YSPTSPS) repeats, is subject to differential phosphorylation during transcription elon-

gation resulting in a distinct, temporally changing phosphorylation pattern, named CTD code. The CTD code determines the factors that associate with the Pol II elongation complex ensuring RNA processing (RNA capping and splicing enzymes) as well as the covalent modification of chromatin (histone-modifying enzymes) [36, 37]. In turn, post-translational modifications of histone tails recruit chromatin remodeling enzymes that are required to enable Pol II to transcribe through chromatin [7]. In this way, the CTD code plays a key role in orchestrating the interplay between transcription and chromatin [38].

Transcription termination

Termination of transcription occurs when the transcription machinery encounters the poly-A site at the 3' end of a gene (reviewed in [39]). The full-length RNA transcript is cleaved and polyadenylated and both processes have been extensively studied [40]. However, the mechanisms underlying transcription termination are still poorly understood. Following termination, Pol II is released from the template DNA and can enter facilitated reinitiation through association with the promoter bound scaffold complex [41].

1.1.2. Structure of the Pol II elongation complex

Crucial for an understanding of the mechanism of transcription is the knowledge of the structure of the transcription machinery. On this account, extensive research effort over the past decades has been concentrated on determining the structure of Pol II in different stages of transcription (reviewed in [27]) and most recently, in complex with key general transcription factors [42, 43, 44, 45, 46]. The first insight into the structure of Pol II was provided by the laboratory of Roger Kornberg in 1991. The overall shape of Pol II was revealed by electron microscopy at 16 Å resolution [47]. Ten years later, near atomic models of the 10 subunit Pol II core [48] and Pol II in complex with a DNA-RNA scaffold [49] were uncovered at 2.8 and 3.3 Å resolution in the same laboratory. During the following years, the picture was rounded off by the elucidation of the structures of the complete 12 subunit Pol II including the Rpb4/7 heterodimeric subcomplex at 4.2 Å resolution [50] and of the complete elongation complex including its interaction with NTPs and TFIIS at 4 Å [51]. In the following, I will describe the structure of Pol II in more detail. Since my study of the Pol II open promoter complex using smFRET and NPS analysis is based on the crystal structure of the complete Pol II elongation complex (EC), I will focus on this structure.

Pol II consists of a 10-subunit core enzyme and a peripheral heterodimer of subunits Rpb4 and Rpb7 that can dissociate from the core [52] (Figure 1.1.2). The Pol II core captures a crab claw-like structure with the two largest subunits Rpb1 and Rpb2 forming opposite sides of a positively charged cleft and remaining subunits being assembled in the periphery (Figure 1.1.2). The Rpb1 side of the cleft forms a mobile clamp, which leads to two different conformations of the free Pol II core, an open and a closed one. However, in the presence of a nucleic acid scaffold, the clamp is stabilized in its closed conformation

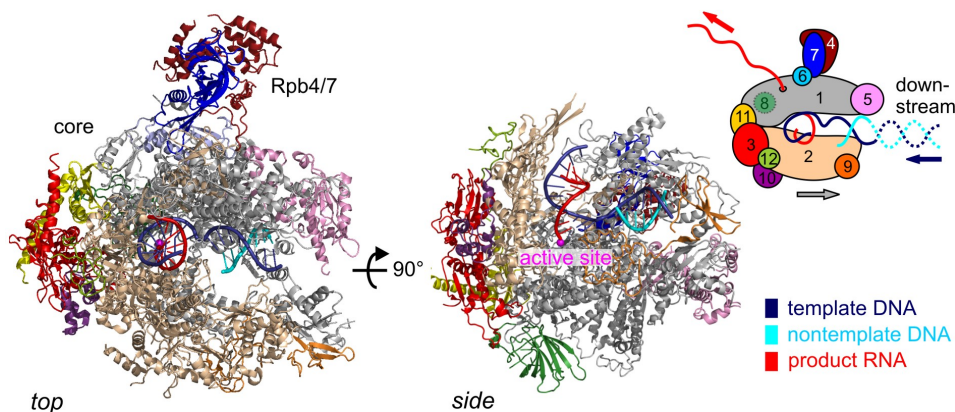


Figure 1.1.2.: Structure of the complete 12-subunit RNA polymerase II elongation complex (pdb ID: 1Y1W, [51]). Two views, the top view (left) and side view (right), are shown with Pol II subunits Rpb1-12 color coded as indicated in the schematic illustration on the right. Template DNA, non-template DNA, and product RNA are shown in blue, cyan, and red, respectively. The active site magnesium ion is shown as magenta sphere. In the schematic illustration (right), the translocation of RNA, DNA and Pol II is indicated by a red, blue and gray arrow, respectively.

[53, 49]. The active site, defined by a persistently bound metal (Mg) ion, is located in a pore at the floor of the cleft. The EC contains a nucleic acid scaffold with an 11 nt melted region, the transcription bubble. The template DNA strand in the melted region binds to the DNA active cleft and is hybridized to about 8 nt of the RNA transcript emerging from the active site. It is aligned such that the 3'-end of the growing RNA is localized next to the active site. The template DNA nucleotide at the active center is numbered +1 and all nucleotides ahead of Pol II that remain to be transcribed are given positive ascending numbers (downstream). The template DNA nucleotide that base-pairs the 3'-end of the growing RNA is numbered -1, and all nucleotides that have been transcribed already are numbered with negative integers (upstream). NTPs enter Pol II via the NTP entry channel, bind the enzymatically active site and interact with template DNA +1. If the NTP is complementary to the template DNA nucleotide, a covalent bond with the growing RNA 3'-end is established under release of a pyrophosphate ion. Subsequently, the DNA-RNA hybrid translocates upstream by one nucleotide to allow for binding of the next NTP (Pol II translocates downstream, see arrows in schematic illustration, Figure 1.1.2). Nascent RNA exits Pol II through the RNA exit channel. In the elongation complex structure, template DNA (+9)-(-10), non-template DNA (+2)-(+9) and 10 nucleotides of nascent RNA, 7 of which base-paired to template DNA (-1)-(-8), could be observed. Due to flexibility, other parts of the used nucleic acid scaffold could not be detected in the x-ray crystallographic structure.

1.1.3. Specific function of initiation factors comprised in a minimal Pol II OC: TFIIB, TBP and TFIIF

In the absence of complete structural information of Pol II initiation complexes, different aspects about the specific function of each GTF during initiation come from numerous structural and functional biochemical studies. I will introduce existing knowledge about the three factors TFIIB, TBP and TFIIF, which are contained in the minimal Pol II open promoter complex studied in chapter 3.

TFIIB has a crucial role in initiation as it is required for the formation of the CC and its transition to the OC and EC. It recruits Pol II to the promoter by binding with its N-terminal zinc ribbon (B-ribbon) domain to the Pol II dock domain [45, 54] and with its C-terminal (B-core) domain comprising two cyclin folds [55] to the promoter DNA, TBP [55] and the Pol II wall [42]. The region connecting the B-ribbon with the B-core forms two elements, the B-linker and B-reader, which are apparently involved in DNA opening [42] and TSS selection [56, 57, 58], respectively. TFIIB is displaced upon EC formation [34]. Recently, the structure of the Pol II-TFIIB complex, in which all except for the C-terminal cyclin fold of TFIIB is visible, was solved and led to a model of the CC [42]. TBP as part of the TFIID complex [59, 60] binds specifically the TATA box core promoter motif, which is in mammals most commonly found at position -31 or -30 relative to the TSS [61, 62]. The crystal structure of a TBP-TATA complex showed that TBP binds with its concave surface to the eight base pairs of the TATA box, bending the DNA by $\sim 80^\circ$ towards its major groove [63]. Beside TBP, TFIID contains 14 TBP-associated factors (TAFs) that interact with gene specific activators, thereby giving TFIID the role of a coactivator.

TFIIF is required for initiation [64], for stable PIC formation [61] and for start site selection [65, 66, 67]. It binds Pol II tightly and is associated to about 50 % of Pol II in yeast cells [68]. Moreover, it prevents nonspecific interaction of Pol II with DNA [69]. The architecture of the Pol II-TFIIF complex was recently investigated by a combination of crosslinking and mass spectrometry and TFIIF was shown to bind to the Rpb2 side of Pol II, in particular to the lobe and protrusion domains [70]. Furthermore, TFIIF can modulate the activity of open complexes by either repressing or stimulating initiation [71]. The response to TFIIF was shown to be dependent on the sequence of template DNA within the melted region, however nothing is known about the underlying molecular mechanism.

1.1.4. Single molecule studies of transcription

Single molecule techniques have greatly advanced our understanding of transcription by RNA polymerases by shedding light on the dynamics and heterogeneity of different aspects of the transcription cycle that escape traditional biochemical bulk methods or x-ray crystallography (recently reviewed in [72, 73]). Due to the complexity of the eukaryotic

Pol II transcription machinery, most single-molecule studies have focused on the bacterial enzyme. Bacterial RNA polymerase (RNAP) consists of 5 subunits (β' , β , 2 α monomers and ω), which are similar in structure and sequence to subunits Rpb1, Rpb2, Rpb3, Rpb11, and Rpb6 of the eukaryotic RNA Pol II. Bacterial and eukaryotic RNA polymerases exhibit striking structural similarities in their overall shape (crab claw-like structure with central DNA-binding cleft) and the relative positions of subunits [74, 75]. Bacterial RNAP only requires one transcription initiation factor, sigma, which tremendously eases initiation studies compared to the eukaryotic system.

Transcription initiation has been the subject of several studies. Atomic force microscopy (AFM) was used to probe the overall shape of open promoter complexes of the *E. coli* RNAP. An architecture was revealed, in which the promoter completely wraps around the polymerase over 270° involving extensive contacts of the upstream DNA with the RNAP surface [76]. Further, two studies explored the early conformational dynamics during abortive transcription initiation by bacterial RNAP using smFRET and magnetic tweezers [10, 11]. Abortive initiation involves cycles of synthesis and release of short RNA transcripts (up to 7 nt) that occur until RNAP produces a longer RNA product enabling promoter escape. Both studies demonstrated that initial transcription proceeds through a scrunching mechanism, in which RNAP remains fixed on the promoter DNA, but pulls downstream DNA into itself, thereby keeping extra DNA inside. In this way, energy of NTP hydrolysis is stored in order to break RNAP free from its tight contact with the promoter later on during promoter escape. Moreover, smFRET using confocal and TIRF microscopy was applied to study the transitioning of *E. coli* RNAP from initiation to elongation and the extent of sigma retention [77, 78]. A FRET dye pair was placed on the sigma factor and on upstream or downstream template DNA and changes in smFRET of actively transcribing complexes were observed. These studies showed that sigma is retained during transcription elongation and hence, sigma release is not required for promoter escape.

Transcription elongation has been extensively studied by single-molecule force spectroscopy using optical tweezers and details of the kinetics of translocation, pausing and backtracking by RNA polymerase could be elucidated. Bacterial RNAP was found to stall at a force of 25 pN, however translocation velocity was largely independent of applied force below that maximal force [79]. Translocation by RNAP was interrupted by stochastic pausing events [79, 80, 81] that were intermediary to irreversible RNAP arrest [80]. Lower nucleotide concentrations were shown to result in a decreased translocation velocity and an increased probability for pausing [82]. Moreover, an investigation of the sequence dependency of pausing suggested that pausing occurs at DNA sequences similar to known regulatory pause sequences [83]. As the resolution of optical tweezers improved, more detailed studies of RNAP pausing became possible and RNAP backtracking was directly observed [84, 85]. In 2007, optical tweezers were for the first time applied to eukaryotic yeast RNA Pol II and three times smaller stalling forces were detected (7 pN) compared

to the bacterial RNAP due to a greater tendency to backtrack [86]. Transcription elongation factor TFIIS, which activates Pol II to cleave the 3'-end of backtracked RNA, was demonstrated to increase the stall force of Pol II by 3-fold suggesting that the low mechanical strength of Pol II is part of the mechanism of transcriptional regulation. In our lab, smFRET in combination with Nano-Positioning System (NPS) analysis was used to study the conformation of regions of the Pol II EC that are not visible in the Pol II EC crystal structure [20, 15]. In this way, the position of the 5'-end of exiting RNA [15], the influence of TFIIB on RNA position [20], and the course of nontemplate and upstream DNA [21] could be revealed.

Transcription termination is the least investigated phase of transcription. Only few studies exist, which probe the importance of mechanical force on termination by the bacterial RNAP using optical tweezers [87, 88] and the kinetics of transcription termination using tethered particle motion assays with immobilized bacterial RNAP [89].

Recently, single molecule studies of transcription were brought to a new level by investigating the interplay between transcribing RNA polymerases and nucleosomes. Optical tweezers were used to observe yeast Pol II during transcription of a DNA template containing a single nucleosome [90]. Pol II was observed to pause in front of the nucleosome and wait for transient unwrapping of the nucleosomal DNA (breathing, see section 1.2.2) in order to invade the nucleosome rather than actively disrupt DNA-histone contacts. Further, by means of AFM the presence of looped transcriptional intermediates were observed providing insights into the histone transfer process during transcription through nucleosomes [91]. A fraction of transcribed nucleosomes was remodeled into hexasomes. Another optical tweezers study used the DNA unzipping technique to determine the position of polymerase and nucleosome upon transcription for a defined time. The bacterial RNAP was used as a model system, even though it does not encounter nucleosome obstacles *in vivo*. RNAP performed backtracking by 10-15 bases when it encountered a nucleosome, however backtracking was reduced in the presence of a trailing polymerase suggesting that the presence of multiple polymerases *in vivo* facilitates transcription through nucleosomes [92].

Even though these more complex *in vitro* experiments combining the transcription process with nucleosomal compaction of DNA describe a more realistic situation, they are still only capable of providing a simplistic mechanistic picture and it remains to be seen how it has to be adapted to describe the complexity within organisms. In recent years, as a result of fast advances in live-cell imaging technologies, it has become possible to apply single molecule fluorescence microscopy to cells and directly observe individual transcription events in single living cells [93, 94, 95]. Studies on the direct detection of nascent mRNA uncovered transcriptional bursting in mammalian cells [95] and bacteria [93], however, single, uncorrelated transcription-initiation events in yeast [96, 97]; further a high cell-to-cell

variability of mRNA production and an inherently probabilistic nature of the transcription process (intrinsic noise). Moreover, transcription of mRNA in yeast was correlated with the diffusion and binding behavior of a corresponding transcription factor, which revealed that transcription initiation of the investigated gene was dependent only on the success of the transcription factor in its search for its particular promoter binding site [97]. Two different approaches were applied to label mRNA. The first labeling approach is single RNA fluorescence *in situ* hybridization (RNA-FISH), in which a fluorescently labeled DNA oligomer probe, complementary to the mRNA of interest, is added to fixed cells and hybridization leads to labeling of the RNA of interest. The second approach is the MS2- or PP7-stem loop technique, which has the advantage that it does not require fixation of cells and can therefore be applied in living cells. In this labeling technique, nascent RNA can be detected by genetically inserting a cassette into the untranslated region (UTR) of the gene of interest, which codes for several repeats of a hairpin binding site of the PP7 or MS2 bacteriophage coat protein. As the cassette is transcribed by Pol II, RNA stem loops form and are bound by the PP7/MS2-fluorescent protein fusion constructs that are constitutively coexpressed and hence serve as fluorescent tag. Such single molecule, single cell investigations will pave the way to a more mechanistic understanding of the dynamics of transcription regulation in the complex cell environment.

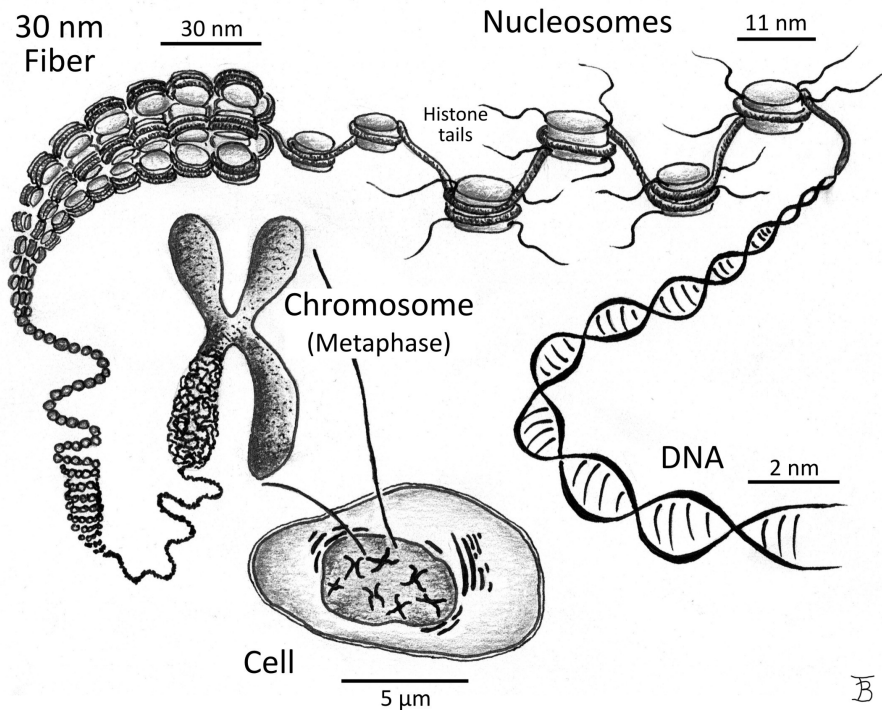


Figure 1.2.1.: Chromatin organization in eukaryotes. The different levels of DNA compaction are presented.

1.2. Chromatin structure and dynamics

1.2.1. Chromatin organization in eukaryotes

In eukaryotic cells, genomic DNA is packaged into a compact nucleoprotein form called chromatin. Chromatin not only allows compaction and organization to make genomic DNA fit into the cell nucleus, but also provides regulatory functions by controlling DNA accessibility. It contains multiple levels of DNA compaction as presented in Figure 1.2.1. The first level of compaction is the 11 nm fiber, which resembles beads on a string [98]. It consists of nucleosomes (beads), the basic packing unit of chromatin, connected by linker DNA (string). In the nucleosome, about 147 bp of DNA are wrapped in 1.65 turns of a flat, left-handed superhelix around a core histone octamer (Figure 1.2.2). Each octamer is composed of two H3-H4 dimers bridged together as a stable tetramer and flanked by two separate H2A-H2B dimers [5, 99]. Histones are rich in basic amino acids, which provides the basis for a stable assembly with negatively charged DNA. The minor groove of the DNA double helix faces inwards and interacts with histones approximately every 10 bp, leading to fourteen contact sites within 147 bp of nucleosomal DNA. Protruding from the nucleosome core particle are the N-terminal histone tails, which are rich in basic

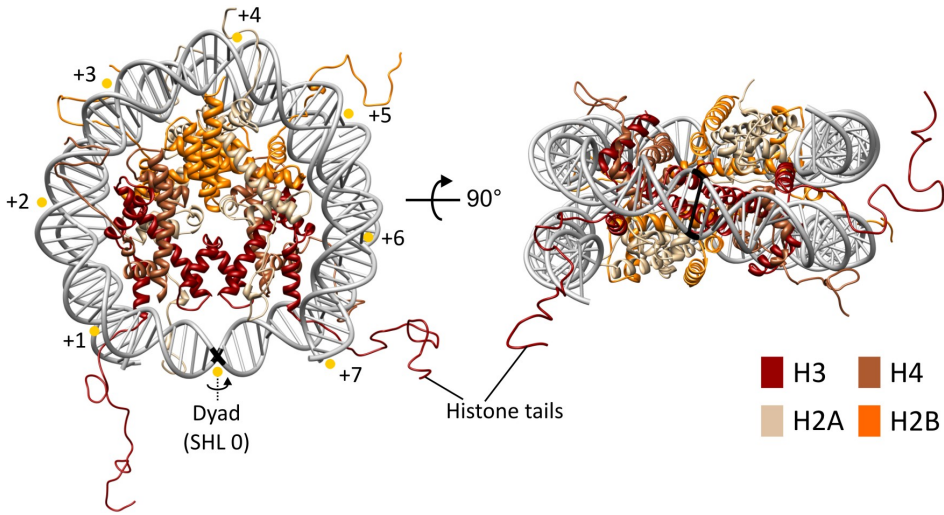


Figure 1.2.2.: Crystal structure of the nucleosome core particle [99]. Nucleosomal DNA (gray) is wrapped around a core histone octamer consisting of a (H3-H4)₂ tetramer (red, brown) and two peripheral H2A-H2B dimers (beige, orange). The dyad (superhelical location 0), the center of nucleosomal DNA through which the pseudo C2 axis of the nucleosome runs, is indicated in black. Superhelical locations (SHL) 1-7 on the first turn of nucleosomal DNA are marked by yellow circles (major grooves). The interaction sites between DNA and histones are situated at SHLs 0.5, 1.5, ..., 6.5 (minor grooves).

amino acids and subject to covalent posttranslational modifications (reviewed in [100]). Beyond the core histones, linker histone H1 can bind internucleosomal linker DNA at the entry/exit site of the nucleosome, thereby creating the chromatosome.

On the next level of chromatin organization, nucleosomal arrays are compacted into 30 nm chromatin fibers. The structure of the 30 nm chromatin fiber remains unknown and discrepancies exist between proposed structural models [101]. During mitosis, chromatin has to be further compacted beyond the 30 nm fiber into the highly condensed chromosomes. In turn, during interphase the structure of chromatin is heterogeneous: while specific genomic regions contain highly condensed chromatin (heterochromatin) inhibiting the transcription machinery, others contain relatively uncondensed chromatin (euchromatin) accessible to factors that bind and act on DNA. However, even euchromatin that mainly consists of uncondensed nucleosomal arrays, can severely impede transcription [102, 103]¹ and, in turn, also heterochromatin exhibits heterogeneity and is plastic. To facilitate transcription of nucleosomal DNA, cells are equipped with a set of specialized multi-protein chromatin remodeling enzymes that use the energy of ATP hydrolysis to modify the structure of chromatin (section 1.2.3).

¹Note that nucleosomes not necessarily impede transcription since Pol II was reported in biochemical and single molecule studies to be able to transcribe through nucleosomes [104, 105, 90, 92].

1.2.2. Nucleosome stability and dynamics

The nucleosome is characterized by 14 interaction sites every ~ 10 bp between the individual DNA minor grooves and the histones [5]. At each site, an arginine is inserted deep into the negatively charged minor groove of the DNA and several hydrogen bonds between the histone main chains and DNA phosphates are established. None of the interactions between DNA and histones is sequence specific, but nonetheless, the stability of a nucleosome, defined as the relative free energy for nucleosome assembly, depends on the sequence of nucleosomal DNA [106, 107] and varies by more than 4 kcal mol^{-1} . This sequence dependency stems from the intrinsic curvature and bendability of certain DNA sequence motives, which have to be arranged such that the sharp bending of DNA within the nucleosome is facilitated. DNA sequences that generate a curved DNA structure favoring nucleosome formation are characterized by 10 bp periodic AA/TT/TA dinucleotides oscillating in phase with each other and out of phase with 10 bp periodic GC dinucleotides. A/T sequences prefer minor groove sites facing the histone octamer and G/C rich sequences favor major groove sites facing outwards. As a consequence of differences in nucleosome stability, DNA sequence also determines nucleosome positioning *in vitro* and *in vivo* [108, 109]. Nucleosome positioning has been studied recently also on the single molecule level using a high throughput single molecule fluorescence technique that involved long DNA curtains [110]. In a selection experiment for DNA sequences that exhibit high affinity for histone octamers, and concomitantly high nucleosome positioning power, the *601-sequence* was identified as exceptionally strong nucleosome positioning sequence [111]. This sequence will be used in all nucleosome experiments presented in this thesis.

A detailed map of the interactions between DNA and histones was constructed in a single molecule study that employed optical tweezers to mechanically unzip single DNA molecules containing a positioned nucleosome [112]. A 5 bp periodicity of interactions was observed to be enveloped by three broad regions of strong interactions. The strongest interaction was found at the dyad and the other two approximately 40 bp away from the dyad at both sides. Unzipping up to the dyad allowed recovery of a canonical nucleosome, however unzipping beyond the dyad led to removal of the histone octamer from the DNA. These results were in consistency with an earlier study probing nucleosome dynamics during unwrapping using optical tweezers [113].

The crystal structure of the nucleosome core particle gives the impression of a static and unchangable assembly. However, nucleosomes hold inherent dynamic properties. Nucleosomal DNA, most of the time buried inside the nucleosome, can be transiently exposed by spontaneous unwrapping. This so-called nucleosome breathing was first revealed in biochemical assays that probed the accessibility of internal sites on the nucleosomal DNA to restriction enzymes [114]. Later, nucleosome breathing was confirmed by single molecule FRET studies of freely diffusing and surface immobilized nucleosomes using confocal and TIRF microscopy, respectively [13, 115]. Recently, acetylation of lysine 56 of histone H3,

present as posttranslational modification *in vivo*, was shown to increase breathing of DNA by 7 fold near the entry/exit site of the nucleosome [116].

In addition to these transient conformational changes, there is recent evidence for alternative nucleosome structures [117]. All existing crystal structures are extremely similar with tightly packed DNA ends near the entry/exit site. However the crystal lattice formation itself selects for such compact state. Many nucleosomes exhibit structural features that are not compatible with crystal formation or that lead to fragile crystals that diffract poorly. This suggests that a structural variability of nucleosomes exists in solution. A recent single molecule FRET study directly proved the existence of an open conformational nucleosomal state even for stable nucleosomes containing the strong positioning sequence 601 at intermediate salt concentration (600 mM NaCl) [118]. This open state is characterized by a partial dissociation of the H2A-H2B dimers from the H3-H4 tetramer near the dyad while remaining bound to the DNA. The possibility that the default structure of a nucleosome in solution does not resemble the crystal structure has to be considered in studies of nucleosome repositioning.

1.2.3. Chromatin remodeling complexes

Chromatin remodeling enzymes are large, mostly multisubunit complexes that use the energy of ATP hydrolysis to modify the structure of chromatin, thereby allowing rapid and regulated access to the underlying genomic DNA. All remodeling complexes are members of the Swi2/Snf2 family of proteins and share as part of their catalytic subunit a conserved DNA-dependent ATPase domain, which belongs to the superfamily 2 (SF2) of DNA translocases [119]. The ATPase domain consists of two parts, DExx (ATPase lobe 1) and HELICc (ATPase lobe 2), that are connected by an insertion [6]. Dependent on unique accessory domains flanking the ATPase domain, remodelers are separated into four distinct families, namely Swi/Snf (mating type switching/sucrose non-fermenting), ISWI (imitation switch), CHD (chromodomain helicase DNA-binding) and INO80 (inositol). The individual families are conserved from yeast to human, although some variation in their detailed protein composition exists. Figure 1.2.3 shows the domain structure of the core ATPase subunit of the four families of chromatin remodelers.

Chromatin remodelers affect the structure of chromatin in different ways and the following remodeling reactions have been identified, both *in vivo* and *in vitro* (reviewed in [6]): sliding or repositioning of histone octamers to a new position on the DNA, partial localized unwrapping of nucleosomal DNA, ejection of histone octamers, ejection of H2A-H2B dimers as well as exchange of H2A-H2B dimers for alternative dimers containing histone variants such as H2A.Z (termed Htz1 in *S. cerevisiae*). The first three remodeling reactions lead to exposure of part of the nucleosomal DNA initially occluded by the histone octamer, whereas the latter two reactions result in an altered composition of the nucleosome itself. Moreover, remodelers have been found to assist in chromatin assembly and

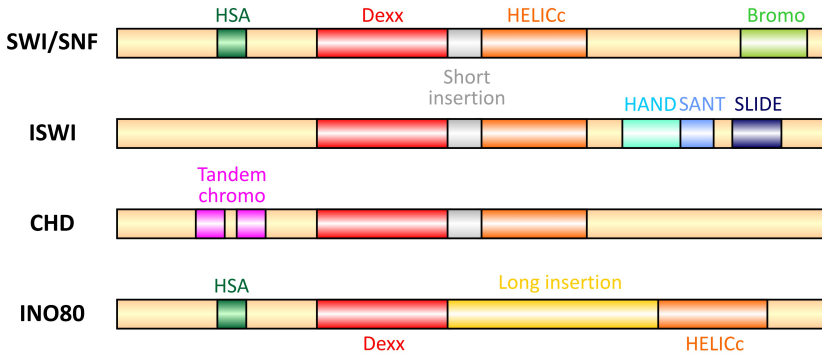


Figure 1.2.3.: Families of chromatin remodelers as defined by the domain structure of their catalytic subunit [6]. All families share a conserved ATPase domain that is split in two parts (DEXX and HELICc) and connected by either a short insertion in the case of SWI/SNF, ISWI and CHD remodelers or by a long insertion in the case of INO80 remodelers. Each family is further defined by a distinct combination of accessory domains flanking the ATPase domain: SWI/SNF remodelers contain a C-terminal bromodomain (light green) and an N-terminal HSA (Helicase-SANT) domain (dark green), ISWI remodelers contain a C-terminal HAND-SANT-SLIDE domain (cyan, light blue, dark blue), CHD remodelers contain N-terminal tandem chromodomains (magenta) and INO80 remodelers contain an N-terminal HSA domain.

induce spacing of nucleosomes within nucleosomal arrays.

Different members of one family can have very different functions with the diversity being mainly imparted by attendant subunits, however some trends can be elaborated: Members of the SWI/SNF family are able to slide and eject nucleosomes, but lack a role in chromatin assembly. Sliding by SWI/SNF remodelers leads to randomized nucleosome positions on a nucleosomal array. In contrast, most ISWI type remodelers were found to slide nucleosomes to create evenly spaced nucleosomal arrays, thereby promoting chromatin assembly and repressing transcription. Also CHD family members were shown to optimize spacing of nucleosomal arrays and assist in chromatin assembly by sliding and ejecting nucleosomes, however most CHD remodelers were shown to have an activating effect on transcription. Only the CHD remodelers Mi-2 and NuRD were shown to have repressive roles. Finally, remodelers of the INO80 family were shown to promote transcription activation and DNA repair, to act as nucleosome spacing enzymes as well as to insert the variant histone H2A.Z by replacing H2A-H2B dimers with H2A.Z-H2B.

1.2.4. The CHD family of chromatin remodelers

Chromatin remodelers of the chromodomain-helicase-DNA-binding (CHD) family consist of one to ten subunits and are characterized by the presence of tandem chromodomains (chromatin organization modifier) in the N-terminal region of the catalytic subunit of the protein [120, 121]. The chromodomain is an evolutionarily conserved sequence motive involved in regulation of chromatin structure and transcriptional regulation. It was first

identified as shared region of the epigenetic repressor proteins heterochromatin protein 1 (HP1) and Polycomb (Pc) [122]. Since then, chromodomains have been found also in ATP-dependent chromatin remodelers, histone acetyltransferases and histone methyltransferases. Members of the CHD family are found from yeast to plants to mammals. The CHD family is divided into three subfamilies according to the presence or absence of additional domains. The first subfamily is defined by a DNA-binding domain located in the C-terminal region and preferentially binding AT-rich sequences [123]. The subfamily consists of *S. cerevisiae* Chd1, which is the only CHD family member in yeast, as well as Chd1 and Chd2 proteins from higher eukaryotes. The second subfamily includes the proteins Chd3 (Mi-2 α) and Chd4 (Mi-2 β) as well as *D. melanogaster* Mi-2 and is defined by the absence of the DNA-binding domain but the presence of an N-terminal PHD (plant homeo domain) Zn-finger-like domain. Members of the second subfamily, Chd3 and Chd4, are part of large complexes consisting of seven proteins that are all associated with transcriptional repression and that were termed NuRD due to their nucleosome remodeling and histone deacetylase activities. The third subfamily contains the proteins Chd5-Chd9 and is characterized by additional functional motives in the C-terminal region, including paired BRK domains, a SANT-like domain, CR domains and a DNA-binding domain.

1.2.5. Chromodomain-helicase-DNA-binding protein 1 - Chd1

Chromodomain-helicase-DNA-binding protein 1 (Chd1) is the founding member of the CHD family of chromatin remodelers and was first identified in mouse as a DNA-binding protein containing in addition a paired chromodomain [124]. Chd1 is conserved from yeast to humans and is the only member of the CHD family of remodelers present in yeast. It is a monomeric remodeler [125] with a molecular weight of 168 kDa and hence significantly smaller than remodelers of the SWI/SNF or INO80 families that consist of up to 15 subunits with molecular weights of around 1-2 MDa [126]. Biochemical studies showed that Chd1 proteins assemble and evenly space nucleosomal arrays as well as slide mono-nucleosomes towards the center of a short DNA fragment [127, 128], similar to ISWI-type remodelers [129, 130, 131]. *In vivo* studies associated Chd1 with active gene transcription by RNA polymerase II, particularly at the level of elongation and termination: Chd1 was shown to localize to regions of decondensed chromatin (interbands) and high transcriptional activity (puffs) in *Drosophila melanogaster* polytene chromosomes [123], to be associated with the transcription elongation factors FACT, Rtf1 and Spt5 at transcribed genes [132, 133, 134] and to participate in transcriptional termination [135]. Moreover, Chd1 was recently linked to transcription initiation since it was shown to be recruited by the Mediator coactivator complex, which controls assembly of the RNA polymerase II preinitiation complex [136]. In *S. cerevisiae*, Chd1 is partially redundant with the action of other spacing enzymes such as Isw1 [137] and a recent genome-wide study demonstrated that the combined action of Isw1 and Chd1 is required to maintain regular nucleosome organization. Moreover, Chd1 was shown to be important for deposition of

the histone variant H3.3 marking active chromatin in *D. melanogaster* [138], and promote nucleosome disassembly in *S. pombe* [139]. In mice it was shown to be essential for open chromatin in order to maintain pluripotency of embryonic stem cells [140].

Chd1 is comprised of a characteristic double chromodomain motive N-terminal to the Snf2 ATPase domain and a C-terminal DNA-binding domain [124, 141] (Figure 1.2.4A). In the following, I will discuss the function and structure of the individual domains of Chd1. Even though the structure of the complete remodeler could not be solved so far, several atomic structures of different portions of Chd1 are available.

Numerous studies have focused on understanding the role of the tandem chromodomains of Chd1: In human, Chd1 chromodomains were found to specifically bind to methylated lysine 4 in histone H3 tails (H3K4me2 and H3K4me3), which might specifically target the function of Chd1 to regions of active transcription marked by H3K4me3 [142, 143, 144]. No interaction of Chd1 was found with unmodified H3K4 or other methylated lysines in H3 tails. In contrast, the tandem chromodomains of *S. cerevisiae* Chd1 do not bind methylated H3K4 peptides [143, 142, 144] and the chromodomains of *D. melanogaster* Chd1 are not required for chromatin localization of Chd1 [145] to actively transcribed regions.

Atomic structures exist for the isolated tandem chromodomains of human and yeast Chd1 (NMR structures, [144]; x-ray crystal structures, [146, 142]) revealing a structural polymorphism that might be responsible for the differing affinity of the two chromodomains to methylated H3K4. Moreover, the crystal structure of a fragment of *S. cerevisiae* Chd1 encompassing the tandem chromodomains as well as the ATPase domain was recently solved and is shown in Figure 1.2.4B [147]. This structure shed light on how the chromodomains negatively regulate the ATPase activity of the Snf2 domain: In the crystal structure, the ATPase motor (red, blue) was observed in an inactive conformation, since the ATPase lobes captured an open conformation that would not allow ATP hydrolysis and since an acidic helix joining both chromodomains (helical linker or chromo-wedge, yellow) occluded the DNA-binding surface of the motor. However, this chromodomain-ATPase interface was shown to be essential for the ability of Chd1 to discriminate between nucleosomes and naked DNA. A model was proposed (Figure 1.2.4B, bottom), in which the specific interaction of a nucleosome with Chd1 stabilizes the helical linker of the chromodomains in an ungated position that enables the ATPase motor to achieve a closed, ATP-hydrolysis competent conformation and therefore allows Chd1 to slide the nucleosome. A recent crosslinking study demonstrated that the ISWI ATPase domain captures such closed, catalytically competent conformation [148].

The DNA-binding domain of Chd1 comprises about 270 residues C-terminal to the Snf2 ATPase domain (residues 1009-1274) and binds to about 20 bp of DNA without sequence specificity but a preference for poly(dA-dT) and poly(dI-dC) sequences [149, 141]. It

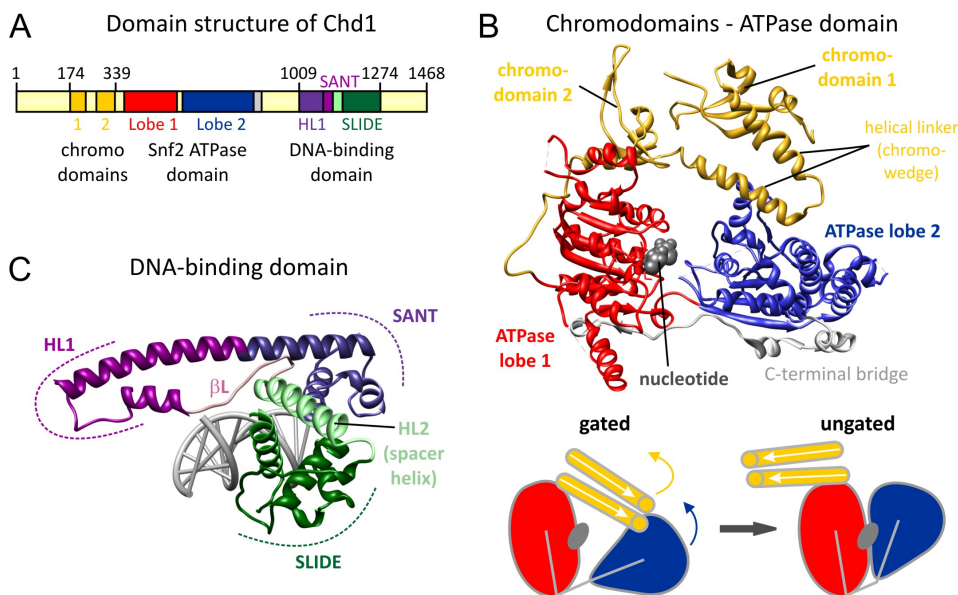


Figure 1.2.4.: Atomic structures of the different Chd1 domains. (A) Domain structure of Chd1. (B) Top: Crystal structure of a portion (residue 142-939) of Chd1 encompassing the tandem chromodomains and the ATPase domain (pdb ID: 3MWY, [147]). The two ATPase lobes shown in red and blue are flanked by the double chromodomain (yellow) and an extended C-terminal bridge (gray). Bottom: A schematic of the crystal structure is shown in the gated conformation with the helical linker of the chromodomains locking the ATPase motor (left) and in the proposed ungated, catalytically competent conformation that is captured upon interaction with a nucleosome and in which the helical linker is displaced and the ATPase cleft has closed. (C) Crystal structure of the DNA-binding domain of Chd1 (residue 1006-1274) in complex with a dodecameric DNA duplex (pdb ID: 3TED, [149, 150]). SANT (purple), SLIDE (dark green), helical linker-1 (magenta, HL1) and -2 (light green, HL2) and β -linker (dark rose, β L) regions are marked. Images were prepared using Chimera [151].

is essential for the stable association with nucleosomes which requires the presence of extranucleosomal DNA, as well as for the nucleosome centering activity of Chd1. However, it was shown that the native DNA-binding domain can be replaced by foreign DNA-binding domains without losing remodeling activity [152]. Furthermore, target sequences recognized by the foreign DNA-binding domains introduced into the extranucleosomal DNA dictated the direction of sliding by Chd1, which suggests that the affinity of the DNA-binding domain of Chd1 for extranucleosomal DNA is the key determinant for the direction of sliding.

In a recent study, the DNA-binding domain of *S. cerevisiae* Chd1 was found to contain a SANT (SWI3, ADA2, N-CoR and TFIIB) and SLIDE (SANT-like ISWI domain) domain analogous to the DNA-binding domain (HAND-SANT-SLIDE domain) of ISWI family remodelers [149, 153]. The fact that both Chd1 and ISWI remodelers act as spacing enzymes and have shared and redundant functions *in vivo* suggests that the common

SANT-SLIDE motive plays an important role in the nucleosome organizing activities. A recently solved crystal structure of the DNA-binding domain of Chd1 in complex with DNA revealed the interaction of surfaces on the SLIDE and SANT domains with a straight DNA duplex consistent with the preference of Chd1 to bind extranucleosomal DNA [150] (Figure 1.2.4C). Unlike the minor groove binding observed for the SANT and SLIDE domains of Isw1 [154], the SANT-SLIDE motive of Chd1 interacts with both major and minor groove of the DNA, thereby establishing a distinct angle between DNA and protein. As a result, the SANT-SLIDE domains of Chd1 span a shorter stretch of DNA than those of Isw1, which could be responsible for the distinct spacing characteristics of both enzymes: While Chd1 establishes spacing with a nucleosome repeat length of about 162 bp, ISWI family remodelers generate nucleosomes spaced 175-200 bp apart [127, 137]. Even though most enzymatic and regulatory functions of Chd1 essential for nucleosome remodeling can be allocated to the three characteristic domains introduced above, it was recently shown that residues in between the ATPase motor and the DNA-binding domain, when altered, greatly reduced the sliding ability of Chd1. These residues therefore appear to play a role in coupling nucleosome stimulated ATP hydrolysis to nucleosome sliding [155].

1.2.6. Mechanisms of nucleosome repositioning and spacing

A nucleosome is held together by 14 histone-DNA contacts that require a total of about 12-14 kcal mol⁻¹ to be broken [156]. This explains the slow rates of translocation of histone octamers along DNA by thermal diffusion and of spontaneous disruption of nucleosomes [157]. Cells have evolved specialized ATP-dependent chromatin remodelers equipped with the capability to overcome these energetic barriers and slide nucleosomes along DNA, thereby providing regulated access to the underlying DNA. However, the mechanism that allows remodelers to couple hydrolysis of ATP to the breakage of histone-DNA contacts without a full disruption of the nucleosome is not well understood. Two distinct models (with slight variations) have been proposed for the remodeling mechanism that both require only a subset of histone-DNA contacts to be broken at a time, preventing disruption of the nucleosome and providing a cost-efficient way of sliding nucleosomes.

The model first proposed, the twist-diffusion model, is based on the observations that the DNA twist within nucleosome core particles can alter [158] and that the Snf2 ATPase domain of remodelers, being bound to one face of the DNA, can move DNA by 1 or 2 bp at a time [159, 160]. In the twist-diffusion model, a twist defect in the DNA is pushed into the nucleosome by the ATPase domain near the nucleosome entry site. This twist then propagates over the octamer surface by one-dimensional diffusion resulting in a corkscrew-like movement of the DNA and a remodeling step size of 1 bp [107, 161, 162] (Figure 1.2.5A,B).

However, numerous recent experiments have questioned remodeling via diffusion of a twist defect in DNA: In a set of biochemical crosslinking studies, the ATPase domain of

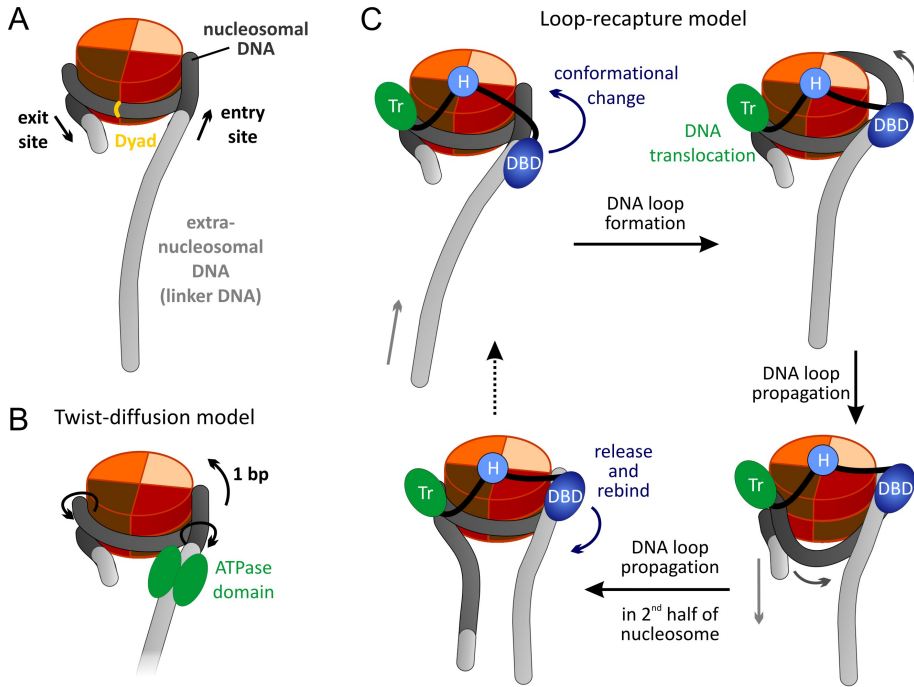


Figure 1.2.5.: Proposed models of nucleosome repositioning mechanism. (A) Schematic illustration of a mono-nucleosome used for remodeling experiments with one short and one long linker DNA. The important elements are labeled. (B) Schematic illustration of the proposed twist-diffusion model [107, 161]. (C) Schematic illustration of the proposed loop-recapture model. The concerted action of a DNA-binding domain (DBD, dark blue) associated with the linker DNA at the entry site and a translocation domain (Tr, green) located near the dyad generates a small DNA loop that propagates over the nucleosome surface. The remodeler is anchored to the histone octamer via the hinge domain (H, light blue). Figure adapted from [6, 163]

ISWI remodelers was shown to interact with nucleosomal DNA around two turns from the dyad at superhelical location 2 (SHL2, on the same gyre as long linker DNA), providing a direct physical connection of ATP hydrolysis with this region of nucleosomes [164, 165, 166, 167]. Consistent with this finding, single-stranded gaps introduced at SHL2 were shown to block sliding of the gapped sites towards the dyad for ISWI remodelers, [166, 165, 168], Chd1 [152] and the catalytic subunit of the SWI/SNF remodeler RSC [169]. Note that a similar gap introduced at SHL-2 on the opposite side of the dyad did not have any effect on Chd1 catalyzed sliding [152]. Moreover, the HAND domain of ISWI remodelers was shown to contact the nucleosomal DNA right at the entry site of the nucleosome and the SLIDE domain was mapped to the extranucleosomal DNA about 30 bp outside the nucleosome core suggesting that this DNA-binding domain senses the length of extranucleosomal DNA. Remodeling by ISWI remodelers was shown to accumulate remodeling products with nucleosomes being preferentially repositioned by a multiple

of 10 bp from their initial position, consistent with the 10 bp periodicity of DNA-histone contacts within the nucleosome [166, 165]. Furthermore, it was shown that ISWI remodelers can slide a nucleosome that has a bead tethered to the nucleosomal DNA near the nucleosomal dyad by 40 - 50 bp [170]. This questions the degree of twist formation during remodeling. As a result of these observations, the nowadays favored loop-recapture (or wave-ratchet-wave) model was proposed for DNA translocation over the histone octamer surface (Figure 1.2.5A,C).

In the loop-recapture model, the remodeler binds the nucleosome core with its ATPase domain (translocation domain) engaging the DNA about 2 turns from the dyad and the C-terminal DNA-binding domain contacts the DNA linker at the nucleosome entry site. In addition, a hinge domain anchors the remodeler to the histone octamer. An ATP-dependent conformational change triggered by the ATPase domain causes the DNA-binding domain to push DNA from the linker on the entry site into the nucleosome thereby creating a DNA loop/wave. The ATPase domain then pumps DNA towards the nucleosome dyad [171] and the DNA loop finally propagates around the nucleosome by one-dimensional diffusion and is released on the nucleosome exit site.

Remodelers of the ISWI family as well as Chd1 share the ability to evenly space nucleosomal arrays. Recently, the structure of two members of the ISWI family, yeast ISW1a and human ACF (ATP-dependent chromatin assembly factor), in complex with a nucleosome were solved by electron microscopy [154, 172] yielding important insights into the mechanism of nucleosome spacing. ISW1a (Δ ATPase) was found to bind two DNA linkers at once and interact with a nucleosome in two different conformations dependent on the extranucleosomal DNA. If the nucleosome contained linker DNA at both sides, both linkers were bound by a single ISW1a, however if the nucleosome contained only one linker, two nucleosomes were bridged by two copies of ISW1a.² In combination with crosslinking data as well as crystal structures of ISW1a(Δ ATPase) alone and bound to DNA, a mechanism for nucleosome spacing was proposed, in which a single ISW1a remodeler uses a dinucleosome substrate and its own physical dimensions as a ruler to set the spacing between two adjacent nucleosomes.

In contrast, Snf2h, the catalytic subunit of human ACF, was shown to bind a single nucleosome as a dimer with both monomers being located at opposite sites around 2 turns from the dyad (SHL +2/-2) [172]. The DNA-binding domain of Snf2h (HAND-SANT-SLIDE) as well as extranucleosomal DNA was not visible in the structure. Binding of the ATPase to SHL +2/-2 as well as the cooperativity of nucleosome binding by the two Snf2h monomers was dependent on their nucleotide state. In combination with biochemical and single-molecule data (see section 1.2.7), a model was proposed in which both monomers continuously sample their linker DNA, yet translocation of the nucleosome occurs mainly towards the longer linker DNA since longer extranucleosomal DNA stimulates the ATPase

²Note that this binding behavior could be artificial and an effect of the experimental conditions during electron microscopy.

activity of the associated Snf2h remodeler. The end product of this competition model is a nucleosome with DNA linkers of similar length or an evenly spaced nucleosomal array. In summary, these results suggest that two remodelers of the same ISWI family, even though they share the same functional domains, use a different remodeling mechanism to evenly space nucleosomes. It remains to be shown if the distinction is real or if both models can be converged into one.

1.2.7. Single molecule studies of nucleosome remodeling

To date, the detailed remodeling mechanisms employed by the different families of remodeling enzymes remains unclear. The application of single molecule techniques will be crucial for a progress in our understanding, since they allow for the direct observation of remodeling of individual nucleosomes in real time. To date, only few single molecule studies of nucleosome remodeling exist (reviewed in [173, 174]). Two studies employed optical and magnetic tweezers to examine the action of RSC and SWI/SNF remodelers on bare DNA and nucleosome templates [175, 176]. The formation of DNA loops was revealed with the loop sizes ranging from 20-1200 bp (100 bp in average). Smaller loops could not clearly be resolved from instrument noise and it remains to be shown whether remodeling involves also or even primarily loops smaller than 20 bp. In addition, for the first time RSC complexes were shown to translocate on DNA with high speeds (200 bp s^{-1}) and for considerable distances (420 bp in average) [176]. In another optical tweezers study, DNA unzipping was applied to examine the location and structure of nucleosomes remodeled by SWI/SNF [177]. Nucleosomes were found to be displaced bidirectionally resulting in a continuous position distribution with a characteristic repositioning distance of 28 bp. Moreover, repositioned nucleosomes remained their canonical structure.

Recently, the first single molecule FRET study was published examining the dynamics of nucleosome remodeling by individual ACF complexes [14]. A double labeled nucleosomal construct was used that allowed for the direct observation of nucleosome translocation by a decrease in FRET efficiency, since one fluorophore was attached to the end of the short DNA linker at the nucleosome exit site (3-N-78 nucleosomes) and the other one to histone H2A which remains at a fixed position during remodeling. ACF was observed to induce gradual translocation of nucleosomes towards the center of the DNA interrupted by well-defined, ATP-dependent pauses. Moreover, bidirectional processive movement by ACF was observed on nucleosomes with long DNA linkers on both sides (78-N-78 nucleosomes). The observed processive back-and-forth movement was explained by the coordinated action of two ACF monomers, that are bound to opposite sides of a nucleosome and oriented for translocation in opposing directions [172, 14].

These first studies indicate the great potential of single molecule techniques applied to nucleosome remodeling, which is waiting to be further exploited in future experiments.

2. Physical basics

In this chapter I give an introduction into the physical basics of the methods used in this thesis. First, I introduce the basics of the optical phenomena fluorescence and fluorescence resonance energy transfer. Then, I discuss the phenomenon of total internal reflection that was exploited in the used TIRF microscope setup, which is introduced afterwards. Subsequently, I give an overview of Hidden Markov Modeling and its application to the analysis of dynamic FRET trajectories and finally, I discuss the basics of Bayesian data analysis that is applied in the global Nano-positioning system used in my studies.

2.1. Fluorescence

2.1.1. Basics

In 1845, John F. W. Herschel reported on the "...superficial colour presented by a homogeneous liquid internally colourless..." [178]. What he had observed was the emission of fluorescence by a solution of quinine sulfate. However, it was only seven years later that George G. Stokes published the first physical description of this optical phenomenon and introduced the term *fluorescence* after the mineral fluorite, which Stokes had found to be able to convert invisible ultra-violet radiation into visible light [179].

Fluorescence describes the spontaneous emission of light accompanied by the transition of a molecule from the first or a higher excited electronic singlet state (S_1 , $S_2 \dots$) to the ground state S_0 [180, chapter 1]. The excited electronic state is usually populated via absorption of energy in the form of light. A *Jabłoński diagram* as shown in Figure 2.1.1A illustrates fluorescence and other possible energy transitions a molecule can undergo upon excitation by light. Initially, the fluorophore is present in its electronic ground state with low vibrational excitations. Through *absorption* of a photon of sufficiently high energy, the molecule is transferred into a higher vibrational energy level of the first excited electronic state S_1 . Since this electronic transition occurs on the time scale of femtoseconds, which is fast compared to the nuclei movement, the nuclear coordinates can be assumed as fixed (*Franck-Condon principle*). Caused by collisions with solvent molecules, the molecule then relaxes within picoseconds to the lowest vibrational state of the first excited electronic state (*internal conversion*, IC).

From there, different radiative and non-radiative processes can occur. The molecule can decay within nanoseconds into a higher vibrational level of the electronic ground state

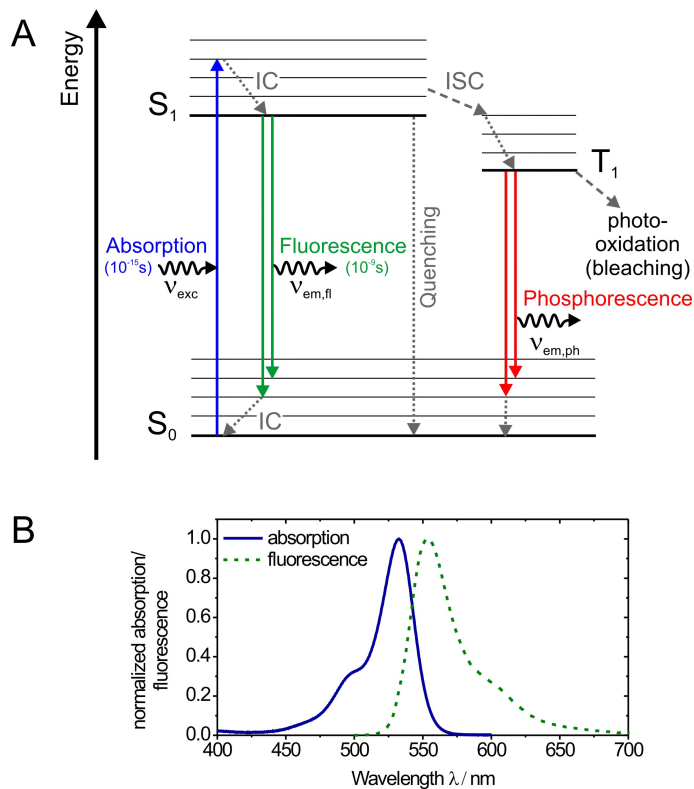


Figure 2.1.1.: Schematic illustration of fluorescence and involved processes. (A) Jablonski diagram as schematic illustration of the possible energetic transitions of a fluorophore upon absorption of light. The energies of the electronic ground state (S_0) and the first electronically excited state (S_1) are shown (thick horizontal lines) together with the associated vibrational states (thin horizontal lines). Absorption (blue), fluorescence (green) and phosphorescence (red) are shown as solid arrows and radiation-less energy transitions are shown by dotted or dashed gray lines (IC, internal conversion; ISC, inter system crossing). ν_{exc} , frequency of excitation photon; $\nu_{em,fl}$, frequency of emitted fluorescence; $\nu_{em,ph}$, frequency of emitted phosphorescence. See main text for description of processes. (B) Absorption (solid blue) and fluorescence emission (dashed green) spectrum of a typical fluorophore, Atto532 (*Atto Tec*, spectrum downloaded from www.atto-tec.de). The emission spectrum presents the mirror image of the absorption spectrum. Further, the stokes shift between absorption and emission is visible.

by spontaneous emission of a photon of frequency $\nu_{em,fl}$ (*fluorescence*) and relax further by internal conversion to the vibrational ground state. Alternatively, the molecule can dissipate its excess energy and return to the electronic ground state by non-radiative processes such as quenching through collision or complex formation with other molecules. Furthermore, spin conversion can occur and the molecule transfers to the first triplet state T_1 (*inter system crossing*, ISC). Since relaxation of the triplet state T_1 back to the singlet ground state is spin forbidden, the T_1 state is relatively long-lived and emission of a photon of frequency $\nu_{em,ph}$ from the T_1 state occurs at a low rate (*phosphorescence*). Another consequence of the longer lifetime of the T_1 state is its increased photochemical reactivity that gives rise to irreversible photo-oxidation and permanent bleaching of the molecule [181].

The phenomenon of fluorescence displays several general characteristics: Fluorescence is about four orders of magnitude slower than vibrational relaxation and hence, emission of fluorescence always occurs from the vibrational ground state of the first excited electronic state. As a result, the fluorescence emission spectrum is independent of the excitation wavelength (*Kasha's rule*). Moreover, the vibrational levels of the electronic ground state S_0 and the excited electronic state S_1 are similarly spaced as a result of the unchanged nuclear geometry (Franck-Condon principle). Therefore, the emission spectrum is usually an approximate mirror image of the absorption spectrum with similar vibrational structures (Figure 2.1.1B). Finally, since absorption and emission occur usually from the vibrational ground state of the respective electronic state but into an excited vibrational state, the emitted photon has a lower energy than the absorbed photon. As a consequence the fluorescence spectrum is shifted to longer wavelengths relative to the absorption spectrum [180, chapter 1] (Figure 2.1.1B). The difference between the maximum wavelength of absorption and emission is called the *Stokes shift* and it is crucial for the application of fluorescence to study single molecules, since it allows for the spectral separation of the emitted photons from the excitation light.

Fluorophores, i.e. molecules that are able to fluoresce, are characterized by three different parameters. The *absorption* or *extinction coefficient* ϵ is a measure for the probability that a photon will be absorbed by the fluorophore. The *quantum yield* Q describes the ratio of the average number of emitted photons to the average number of absorbed photons. It can be expressed in terms of rates as

$$Q = \frac{\Gamma}{\Gamma + k_{nr}}, \quad (2.1)$$

where Γ is the rate to emit a photon and k_{nr} is the rate of the total of non-radiative processes [180, chapter 1]. The third parameter is the *fluorescence lifetime* τ , which is the average time the molecule spends in the excited state and is the inverse of the fluorescence decay rate Γ .

2.1.2. Fluorescence anisotropy

Absorption of a photon involves the coupling of the absorption transition dipole moment of the fluorophore with the electric field of the incident photon. The transition dipole moment of a fluorophore has a particular orientation relative to its molecular axes. Therefore, the probability of electromagnetic coupling between the electric vector of the incident photon and the transition dipole of the fluorophore depends on the relative angle between them and is highest, when both moments are oriented in parallel at the instant of absorption (*photoselection*). In turn, the polarization of the electric field of the emitted photon is determined by the orientation of the emission transition dipole moment of the fluorophore at the instant of photon emission. In solution experiments, excited fluorophores, that are mostly oriented parallel to the electric field of light during excitation, lose their preferential orientation with time by orientational diffusion before emitting a photon. Hence, the extent of polarization of the emitted light decays in time. The correlation between the polarizations of the incident and the emitted photons can be quantified by the steady-state *fluorescence anisotropy* r , given by

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}}, \quad (2.2)$$

where I_{\parallel} and I_{\perp} denote the intensities of fluorescence light with polarization parallel and perpendicular to the polarization of the excitation light, respectively. In biological applications, fluorophores are often attached to large macromolecules via flexible linkers and therefore exhibit constrained orientational and spatial mobility, which leads to an increased anisotropy. Moreover, the time dependent decay of anisotropy is bi-exponential due to the different time scale of rotational motion of the fluorophore and the large macromolecule, to which it is attached.

2.1.3. Fluorescence Resonance Energy Transfer (FRET)

When two fluorophores are in close proximity (usually $< 10 \text{ nm}$), their electronic systems can interact with each other and through dipolar coupling, excited state energy of one fluorophore (donor) can be transferred radiation-less to the other fluorophore (acceptor) in a photophysical process called *Fluorescence resonance energy transfer* (FRET). Theodor Förster was the first to give a quantum mechanical description of the FRET process, which is therefore often also referred to as Förster resonance energy transfer [182, 183]. The Förster theory considers fluorophores as infinitely small point emitters and this approximation only holds well for inter-fluorophore distances larger than the size of the electronic system of the fluorophores. The FRET process is illustrated in the Jabłoński diagram in Figure 2.1.2A. The donor fluorophore is initially excited by absorption of a photon of frequency ν_{exc} followed by a transfer of the excitation energy through long range dipol-dipol interaction without emission of a photon to the acceptor fluorophore. Effectively, the donor is quenched into the ground state and the acceptor is simultane-

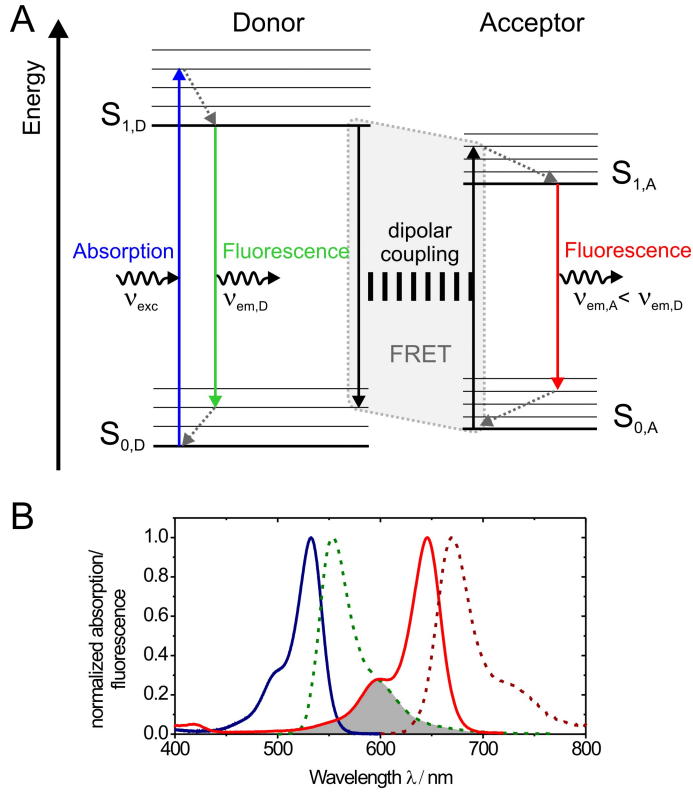


Figure 2.1.2.: Schematic illustration of FRET. (A) Jablonski diagram of the FRET process. The electronic and vibrational states of a donor fluorophore and an acceptor fluorophore are shown. Upon absorption of a photon by the donor fluorophore, its excited state energy can be transferred to the acceptor fluorophore through dipolar coupling (lines between both term systems). In this way, the donor fluorescence is quenched and instead the acceptor emits a photon of frequency $\nu_{em,A}$ during relaxation to its ground state. (B) Fluorescence emission (dashed green) and absorption spectrum (solid blue) of the donor dye are shown together with absorption (solid red) and fluorescence emission spectrum (dashed dark red) of the acceptor fluorophore. The spectral overlap of the donor emission spectrum and the acceptor absorption spectrum is shaded in gray.

ously excited. Finally, the acceptor decays to its ground state by emitting a photon of frequency $\nu_{em,A}$. Since energy has to be conserved during the FRET process, an overlap of the donor emission and the acceptor absorption spectrum is required (Figure 2.1.2B).

The energy transfer between two fluorophores can be expressed by a characteristic rate k_{FRET} , which is given by

$$k_{\text{FRET}}(r) = k_D \left(\frac{R_0}{r} \right)^6, \quad (2.3)$$

where r is the distance between the donor and acceptor, R_0 is the characteristic Förster

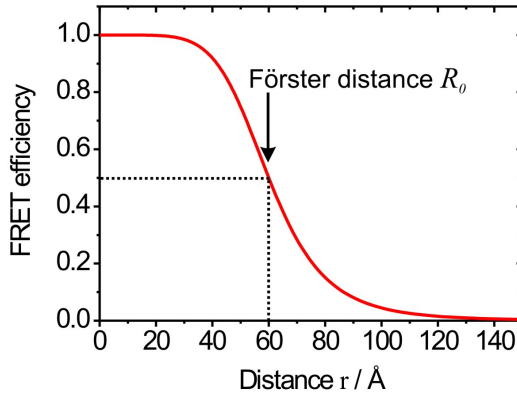


Figure 2.1.3.: Dependency of the FRET efficiency on the donor-acceptor distance r calculated from equation 2.4 with a Förster distance of $R_0 = 60 \text{ Å}$.

distance and k_D is the rate of fluorescence emission by the donor in the absence of an acceptor, which is inversely related to the fluorescence lifetime of the donor τ_D by $k_D = 1/\tau_D$.

Since FRET is competing with other processes that depopulate the excited electronic state S_1 of the donor, the efficiency of energy transfer E_{FRET} can be expressed as ratio of the rate of FRET k_{FRET} to the sum of rates of all competing processes:

$$E_{\text{FRET}} = \frac{k_{\text{FRET}}}{k_D + k_{\text{FRET}}} = \frac{1}{1 + (r/R)^6}. \quad (2.4)$$

Equation 2.4 shows that the efficiency of energy transfer E_{FRET} is strongly dependent on the distance r between donor and acceptor dye. Therefore, FRET is often used as a molecular ruler to measure distances between different sites on macromolecules, as first proposed by Lubert Stryer [184]. This application is facilitated by the fact that FRET is valid in the distance range between approximately 10 to 100 Å, a length scale comparable to the size of biological macromolecules.

2.1.4. Förster distance

The Förster distance, also called Förster radius, is a characteristic parameter for every donor-acceptor pair and is defined as the distance between donor and acceptor that yields a FRET efficiency of 50 % (Figure 2.1.3). It is given (in Å) by

$$R_0 = \left(8.79 \cdot 10^{-5} \cdot \frac{\text{Å}^6 \text{ cm M}}{\text{nm}^4} \cdot \frac{Q_D \cdot \kappa^2 \cdot J(\lambda)}{N_A \cdot n^4} \right)^{\frac{1}{6}}, \quad (2.5)$$

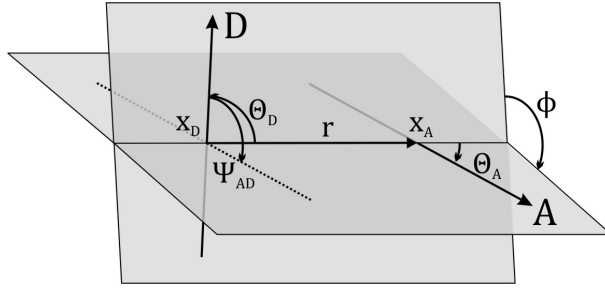


Figure 2.1.4.: Relative orientation of the transition moments of a donor-acceptor fluorophore pair.

where Q_D is the quantum yield of the donor, κ^2 is the orientation factor determined by the relative orientation of the transition moments of both fluorophores, $J(\lambda)$ is the spectral overlap between the donor emission and the acceptor absorption spectrum, n is the refractive index of the medium surrounding the fluorophores and N_A is the Avogadro constant (mol^{-1}).

In order to derive the distance r between a given donor-acceptor pair from the measured FRET efficiency, the respective Förster distance R_0 has to be calculated, which in turn requires that all parameters determining R_0 are known. The refractive index n is usually assumed to be close to that of water ($n = 1.33$). The parameters $J(\lambda)$, κ^2 , Q_D are distinct for every donor-acceptor dye pair and strongly depend on the molecular environment of the dyes, which is usually the surface of a particular biomolecule of interest to which they are attached via flexible linkers. Therefore, these parameters have to be determined for every acceptor-donor dye pair individually.

The quantum yield of the donor is determined by comparison with standard fluorophores such as Rhodamine 101. The spectral overlap $J(\lambda)$ is calculated from the spectral density of the donor emission fluorescence intensity, $f_D(\lambda)$, and the absorption spectrum of the acceptor as a function of the excitation wavelength, $\epsilon_A(\lambda)$, by

$$J = \frac{\int_0^\infty d\lambda f_D(\lambda) \epsilon_A(\lambda) \lambda^4}{\int_0^\infty d\lambda f_D(\lambda)}. \quad (2.6)$$

For a given relative orientation of the transition dipole moments of acceptor and donor fluorophore, the orientation factor κ^2 can be expressed as

$$\kappa^2 = (\cos \psi_{AD} - 3 \cos \theta_D \cos \theta_A)^2. \quad (2.7)$$

Here, ψ_{AD} is the angle between the donor emission transition dipole moment and the acceptor absorption transition dipole moment and θ_D and θ_A are the angles between these dipoles and the line connecting donor and acceptor (Fig. 2.1.4).

The values of κ^2 can range from 0 (perpendicular dipole moments) to 4 (parallel dipole moments). If a random orientation of donor and acceptor dipole moments can be assumed due to rotational diffusion prior to energy transfer, κ^2 is equal to 2/3.

In general, the orientation factor κ^2 is a critical parameter in the quantitative determination of distances from FRET efficiencies, since it cannot be measured easily, but assumption of a wrong value for κ^2 results in serious distance errors [185]. However, it is possible to set limits on the range of κ^2 by determining the steady-state anisotropies of donor and acceptor dye molecules, which is used in the global Nano-Positioning System analysis applied in this thesis (section 2.5).

2.1.5. Site-specific fluorescent labeling of proteins

Most biological applications of FRET require the efficient attachment of a fluorophore to a specific site of a protein. For the studies presented in this thesis, site-specific protein labeling is absolutely essential.

Generally, labeling is achieved by using fluorophore derivatives that can specifically react with a functional group present in the biomolecule. In order to preserve the biological function of the protein, direct attachment of the fluorophore is avoided and instead a flexible linker, usually composed of a chain of multiple methylene groups, is used to tether the fluorophore to the protein.

Two widely used traditional protein labeling strategies are based on the naturally occurring amino acids cysteine and lysine, which have suitable functional groups in their side chain. Cysteine contains a thiol group that can specifically react with a maleimide derivative of a fluorophore and lysine holds an amino group, which reacts with an N-succinimidyl ester derivative of a fluorophore to a stable amide bond. However, lysine is a very common amino acid in proteins and hence, only unspecific labeling can be achieved by this approach. Labeling via cysteine-specific chemistry is well suited for proteins that have only one accessible cysteine. Yet, many proteins contain more than one cysteine residue and as a consequence, prior to labeling, all but one cysteines need to be removed by site-directed mutagenesis (mostly by replacing cysteine by serine, which is similar in size and structure). This approach is applicable to a number of proteins (such as TBP or Chd1, used as labeled single cysteine mutants in this work), however, such point mutations oftentimes negatively influence and modify the function, structure and stability of a protein and in this case different strategies are required.

In recent years, numerous alternative labeling methods have been developed presenting a substantial progress compared to cysteine-specific chemistry. A common feature of the new approaches is the use of genetically encoded residues or small peptide tags that are introduced into the protein and that can be specifically labeled post-translationally in a bio-orthogonal way. One new approach is the Sfp phosphopantetheinyl transferase-catalyzed labeling, which requires the insertion of an 11-amino acid peptide tag, called ybbR-tag, into an unstructured region of the protein [186, 187]. Upon expression of the

protein containing the tag, *Bacillus subtilis* Sfp1 can transfer a fluorophore from coenzyme A to a specific serine residue in the ybbR-tag during formation of a phosphopantetheinyl linker. This strategy was used in the present work to label transcription factor IIB, as presented in chapter 3. A similar strategy is based on the incorporation of a genetically encoded 6-amino acid peptide tag into a protein [188]. The peptide tag contains a cysteine that is converted *in vivo* into an aldehyde by co-expressed formyl-glycine-generating enzyme (FGE) [189]. The aldehyde group then serves as an exclusive target for mild and highly efficient ($\sim 100\%$) labeling with a hydrazide derivative of the fluorophore of choice. Another method, which unlike the previous ones can only terminally label proteins, is the intein-mediated protein ligation (IPL), which requires the N- or C-terminal fusion of the protein of choice to an intein protein. The intein tag can self-catalyze its excision from the protein as well as the ligation of the newly created free protein terminus to a peptide that is supplied in solution and labeled with the desired fluorophore. The highlight of recent advances was the development of a method to encode unnatural amino acids with diverse physicochemical and biological properties genetically in bacteria, yeast and mammalian cells [190, 191]. The amber nonsense codon (UAG) was used to specify a novel amino acid and an orthogonal tRNA:aminoacyl-tRNA-synthetase pair was taken from archaea and modified such that the aminoacyl-tRNA-synthetase could acylate the orthogonal tRNA with the unnatural amino acid but not with endogenous amino acids. Using this method, an unnatural amino acid with the desired chemical functionality can be introduced at any site into the protein and subsequently bound to a fluorophore in a post-translational reaction.

2.1.6. Single molecule fluorescence

The first optical single molecule detection was demonstrated in 1989 by W.E. Moerner and L. Kador, who observed the absorption spectrum of single pentacene molecules embedded in a host crystal of p-terphenyl at liquid helium temperatures [192]. One year later, M. Orrit and J. Bernard used the same system to detect single pentacene molecules by their fluorescence excitation spectrum [193]. However, for the biological application of single molecule fluorescence, the ability to observe fluorophores at room temperature and in aqueous conditions is absolutely required. This was first achieved in 1995 by the use of total internal reflection fluorescence (TIRF, see section 2.2) microscopy [194]. One year later, single molecule FRET (smFRET) was introduced by Taekjip Ha and coworkers [8]. Since then, single molecule fluorescence techniques have revolutionized biological research by giving mechanistic insights into elementary molecular processes previously unattainable [9].

Single molecule approaches provide a number of major advantages compared to ensemble techniques. Oftentimes, samples are heterogeneous either in a static way due to the coexistence of different subpopulations of molecules or because of the dynamic switching of molecules of an ensemble between different states in a stochastic, asynchronous man-

ner. In both cases, only average values of the experimental observables can be obtained by ensemble methods and dynamics as well as heterogeneity are obscured. In contrast, single molecule approaches allow to distinguish between the individual behavior of single molecules and hence subpopulations can be easily resolved. In addition, due to the high time resolution of single molecule fluorescence techniques, structural dynamics of individual molecules can be observed in real time providing a dynamic view of complex biological processes.

The major step towards the detection of the fluorescence signal from single fluorophores was the efficient suppression of background noise by a reduction of the detection volume. This can be achieved in two major ways: Total internal reflection fluorescence microscopy uses an evanescent wave for fluorescence excitation thereby reducing the detection volume to an ~ 100 nm thin layer on the surface of the glass slide. For application of TIRF microscopy, the sample has to be immobilized on the surface of the measurement chamber and all molecules in one field-of-view can be observed simultaneously over an extended period of time. This method was applied throughout this thesis and is discussed in more detail in section 2.2. The other way to decrease the detection volume is realized in a confocal microscope, where the sample is excited with focused laser light and the excitation volume is decreased by the use of a pinhole ensuring elimination of out-of-focus light. In this way, a single spot in solution is detected, which has to be scanned in order to construct an image. In confocal microscopy, molecules are freely diffusing through solution and can therefore only be observed during the instant they pass through the detection volume.

Fluorophores used for smFRET need to meet a number of requirements. First, in order to achieve a high signal-to-noise ratio, fluorophores should be bright and yield a high photon count rate. Moreover, since the observation of a single molecule over an extended period of time is possible, fluorophores should last long (high photostability, slow photobleaching) and not show any temporal fluctuations of fluorescence intensity such as photoblinking. The fast progress in the field of single molecule fluorescence in recent years has stimulated the development of new fluorophores and measurement conditions (oxygen scavenging system, see section 4.6.13) that meet these stringent demands and will further improve single molecule detection.

2.2. Total internal reflection fluorescence microscopy (TIRFM)

Total internal reflection fluorescence microscopy (TIRFM) is a widely used method to detect fluorescence and FRET on the single molecule level. It was first developed for cell imaging [195], but with improved detection technologies, has also become applicable to single molecule experiments [194, 8]. TIRFM makes use of the unique properties of an evanescent wave that is created, when light is not refracted (Figure 2.2.1A, blue) but totally internally reflected at an optical interface. Total internal reflection occurs, when light strikes the interface between a media of higher refractive index n_1 (e.g. glass, $n =$

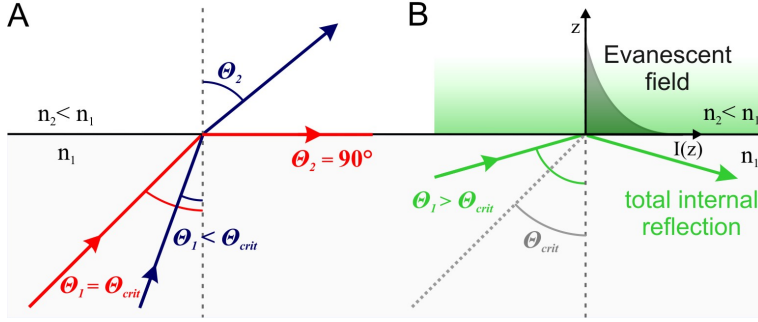


Figure 2.2.1.: Optical geometry for total internal reflection. (A) For an incident angle $\theta_1 < \theta_{crit}$, the light beam is refracted on the interface between media 1 with higher refractive index n_1 and media 2 with lower refractive index n_2 and penetrates into media 2. At the critical angle θ_{crit} , the light beam refracts at an angle of 90° and propagates along the interface. (B) For incident angles $\theta_1 > \theta_{crit}$, refraction does no longer occur and instead the incident light beam is totally internally reflected at the interface. Thereby, an evanescent field is created that decays exponentially along the z -axis.

1.5) and a media of lower refractive index $n_2 < n_1$ (e.g. water, $n = 1.3$) at an incident angle exceeding the critical angle θ_{crit} (Figure 2.2.1A,B).

Snell's law describes the relation between the refractive indices n_1 and n_2 and the angles θ_1 and θ_2 of the incident and the emergent light beam, respectively, as

$$n_1 \cdot \sin \theta_1 = n_2 \cdot \sin \theta_2. \quad (2.8)$$

From this relation, the critical angle θ_{crit} can be derived as

$$\theta_{crit} = \arcsin \left(\frac{n_2}{n_1} \right). \quad (2.9)$$

When total internal reflection occurs, the light intensity penetrates a short distance perpendicular to the interface into the media with lower refractive index (water) and a so-called evanescent wave is generated. This evanescent wave is a standing wave with the same frequency as the incident light that has its maximum intensity at the interface (at the surface of the glass slide) and exponentially decays with increasing distance in the z -direction (Figure 2.2.1B), as described by

$$E(z) = E(0) \cdot e^{-z/d}. \quad (2.10)$$

$E(0)$ is the energy of the evanescent wave at the interface and $E(z)$ is the energy at a distance z perpendicular to the interface. The penetration depth d of the evanescent wave is dependent on the wavelength of the incident light beam λ , the angle of incidence θ_1 and the refractive indices of both media, n_1 and n_2 , and is given by

$$d = \frac{\lambda}{4\pi} \left(n_1^2 \sin^2 \theta_1 - n_2^2 \right)^{-1/2}. \quad (2.11)$$

2 Physical basics

As a consequence of this exponentially decaying evanescent field, only fluorophores located within ~ 100 nm from the glass surface are excited, as opposed to molecules freely diffusing in the bulk solution at distances exceeding 100 nm. In this way, background signal is efficiently suppressed and the signal-to-noise ratio is increased such that single molecules can be detected.

2.2.1. Experimental setup

TIRF microscopy is most commonly realized in one of two different ways. In *objective-type* TIRFM, laser light is focused off-axis onto the back focal plane of an objective, which causes the light to be incident on the glass-water interface at an angle above the critical angle. Both emission and excitation light pass through the same objective and hence, the excitation light is always aligned with the collection optics. A dichroic filter spectrally separates excitation from emission. In *prism-type* TIRFM, which is used in this work, the excitation and emission pathways are separated. The excitation light is guided through a prism attached to the backside of the quartz glass slide, to which the sample is immobilized. The light strikes the water/quartz glass interface at an angle above the critical angle (here 74°) such that it is totally internally reflected away from the objective. Through the objective opposite to the prism, the fluorescence emission is collected, filtered to remove unwanted light and imaged on a low-noise EMCCD camera. The advantage of objective-type TIRFM is its suitability for life cell experiments that are usually cultivated in open dishes. In turn, prism-type TIRFM offers a better signal-to-noise ratio and therefore is preferentially used for *in vitro* single molecule experiments.

All smFRET experiments presented in this thesis were performed on a prism-type TIRF microscope homebuilt by a former PhD student of our laboratory, Robert Lewis. I will give a brief presentation of the TIRFM setup, however an elaborate description can be found in the PhD thesis of R. Lewis [196].

The TIRFM setup consisted of three different excitation sources: a frequency-doubled Nd:YAG laser (532 nm, *Spectra-Physics*) for the excitation of donor fluorophores, a diode laser (637 nm, CUBE 635-25C, *Coherent*) for the direct excitation of acceptor fluorophores as well as a diode pumped frequency doubled solid state laser (491 nm, *Cobolt calypso*) for bleaching background fluorescence prior to the actual measurement. All laser beams were collimated using individual telescopes, combined spatially by dichroic mirrors and guided through an acousto-optic tunable fiber (AOTF, *Opto-Electronic*). The AOTF allowed for selection of the laser wavelength used for excitation, control of the laser intensity as well as control of the duration of excitation. After passing the AOTF, the excitation light was focused onto the prism at an incident angle of approximately 74° in order to achieve total internal reflection. The prism (*Thorlabs*), coated with immersion oil, was mounted on the quartz glass slide of the measurement chamber (see subsection 2.2.2), to which the fluorescent probe was immobilized. The illumination area was about $70 \times 35 \mu\text{m}^2$ with a power

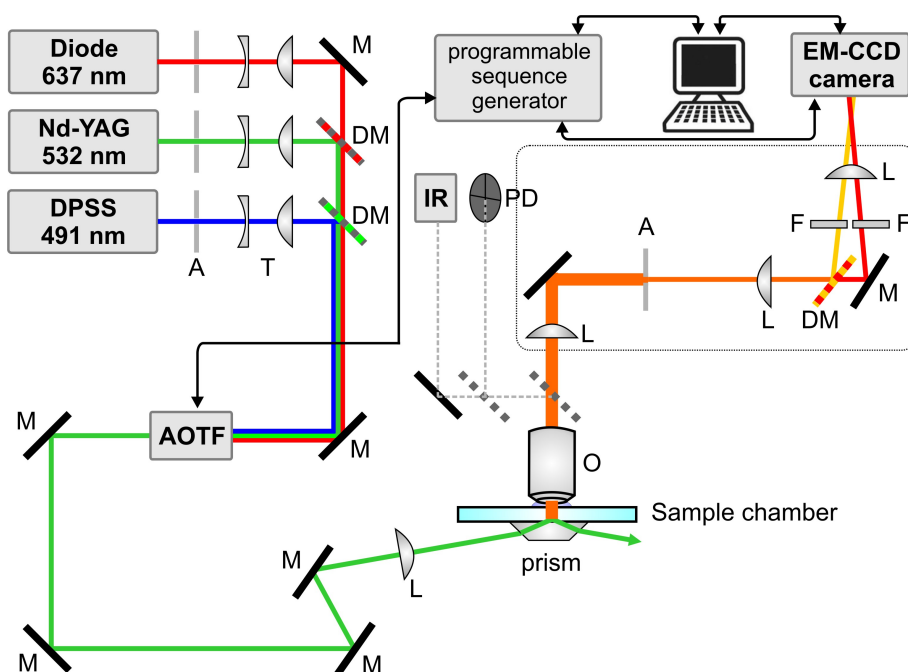


Figure 2.2.2.: Schematic illustration of the experimental prism-type TIRFM setup (adapted from [196]). M, mirror; DM, dichroic mirror; L, lens; AOTF, acousto-optic tunable filter; A, aperture; F, emission filter; O, objective; PD, 4-segment photo diode; T, telescope.

density of $1.5 \mu\text{W}/\mu\text{m}^2$. Upon excitation of the sample, fluorescence intensity of donor and acceptor dye was collected through a water-immersion objective (Plan Apo 60X, NA 1.2, *Nikon*) and directed to an intensified, back-illuminated EMCCD camera (iXon DU-897E-CS0-BV, *Andor*). Two different detection channels were created by blocking one half of the EMCCD-chip using an aperture in the image plane and spatially separating donor and acceptor fluorescence by using a dichroic mirror (*Chroma* 645DCXR). Additionally, emission filters in detection path of both channels isolated emission wavelengths from 550-610 nm (green channel, *Omega Optical* 3RD550-610) and from 660-760 nm (red channel, *Chroma* HQ710-100M). Due to a 120 times magnification through a lens assembly, one pixel corresponded to 133×133 nm.

During an experiment, the observation area was scanned by moving the measurement chamber with a micrometer stage (*Thorlabs*) in x- and y-direction (not shown in Figure 2.2.2). The z position of the sample chamber was controlled by a piezo stage (not shown in Figure 2.2.2), which was embedded in a focus-hold system [197]. This system consisted of an IR laser that was focused through the objective onto the measurement chamber. The back-reflected beam returned through the objective, was separated from the emission light by reflection on a dichroic mirror and was imaged onto a four-segment photo diode incorporated into a feedback loop with the piezo stage. When molecules

on the surface of the measurement were in focus, the IR laser stroke the center of the photo diode. Deflection from the center induced the piezo stage to move the chamber in z-direction until the focus was reestablished. A syringe pump (*Harvard Apparatus*, PHD 2000 Infuse/Withdraw) was used to load sample or exchange buffer during the course of an experiment.

Due to the AOTF, different excitation schemes were possible. Oftentimes, the sample was excited solely with the green laser and movies with 400 frames and an integration time per frame of 100-33 ms were acquired. Alternatively, green and red excitation, corresponding to FRET excitation and direct acceptor excitation, could be alternated with a frequency of 10-100 Hz in so-called ALEX (alternating laser excitation) measurements [198]. This alternation is particularly important when dynamics in the FRET time trajectories are observed. Such dynamics can be caused by either conformational changes within the investigated molecules or by fluctuations in acceptor brightness and acceptor blinking. ALEX allows for the discrimination between these two possible causes and prevents misinterpretation of dynamic FRET trajectories. Most smFRET experiments for the NPS analysis of Pol II open complexes (chapter 3) and Chd1-nucleosome complexes (chapter 4.4) were performed with a constant green excitation of the samples. However, all experiments investigating the nucleosome dynamics during binding and remodeling of Chd1 (chapter 4.3) were carried out in the ALEX mode.

2.2.2. Preparation of flow chamber for smFRET measurements

PEG-passivated flow chambers were constructed in a similar procedure as published previously [199]. A detailed description of the chamber design and preparation can be found in the dissertations of two former PhD students, R. Lewis and A. Andrecka [200, 196].

Briefly, quartz glass slides (*Baumbach*) and coverslips (*Thermo Scientific*) were thoroughly cleaned by boiling first in detergent (*Hellmanex II*, *Hellma*) and after that in ultra-pure water (*Millipore Simplicity 185*). Subsequently, slides were rinsed with ultra-pure water and dried in nitrogen flow. Remaining dust and other particles on the surface of the slides were oxidized using a butane gas torch. In the next step, quartz slides were silanized by incubation in a 2% (vol/vol) solution of (3-aminopropyl)-triethoxysilane (*Sigma-Aldrich*) in acetone for 15 min at room temperature, followed by rinsing with ultra-pure water and blow-drying in the nitrogen flow. During this silanization, the quartz glass surface was functionalized with amino groups, which allowed for covalent attachment of long-chain PEG molecules, functionalized with amino-reactive succinimidyl-groups, in the next step. For PEGylation, a solution of mPEG-succinimidyl-propionate (400 mg/ml, MW 5 kDa, *Laysan Bio, Inc.*) and biotin-PEG-N-hydroxysuccinimid (15 mg/ml, MW 3.4 kDa, *Laysan Bio, Inc.*) in carbonate-bicarbonate buffer at pH 9.4 was prepared and applied to the quartz glass slides for 1 hour at room temperature. To remove excess PEGylation reagent, the quartz slides were again rinsed thoroughly with ultra-pure water and dried in nitrogen

2.2 Total internal reflection fluorescence microscopy (TIRFM)

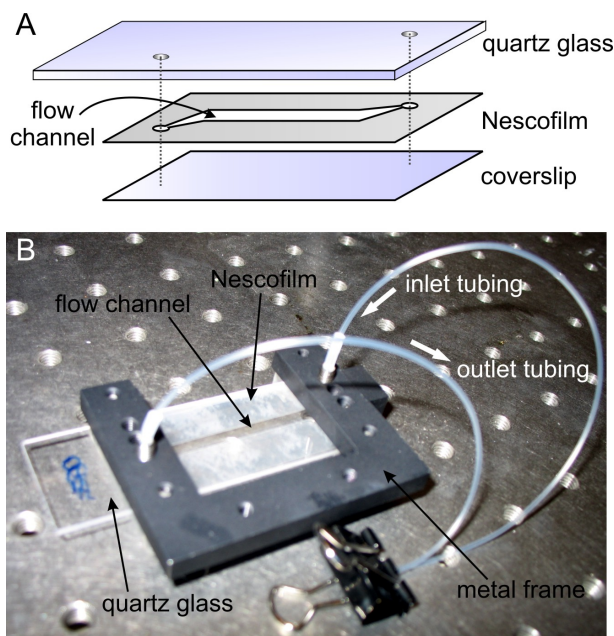


Figure 2.2.3.: Flow chamber design for smFRET measurements on homebuilt TIRF microscope setup. (A) Schematic illustration of assembly of quartz glass, Nescofilm and coverslip to create the flow channel, into which fluorescent sample can be injected. The quartz glass slide is coated at the bottom with PEG/PEG-biotin such that in the final chamber it is accessible to the sample for immobilization. (B) Picture of the complete flow chamber. The main components are labeled. During TIRF measurements, a prism attached to a metal holder is mounted on top of the chamber and fixed with two screws (see screw threads on the upper front part of metal frame).

flow. Finally, each PEGylated quartz slide was assembled together with a pre-cut sealing film (*Nescofilm*) and a coverslip such that the Nescofilm matched the holes in the quartz slide (Figure 2.2.3A). The assembly was then heated to 130 °C for about one minute to allow for the thermoplastic Nescofilm to seal the flow channel created between the holes in the quartz slide. During heating, the assembly of slides and film was weighted down to prevent the formation of air bubbles in the film. In the last step, the flow channel assembly was mounted into a customized holder and inlet and outlet tubings (ID: 0.58 mm, OD: 0.96 mm) were screwed onto the holes of the quartz glass slide (Figure 2.2.3B).

The PEG/biotin-PEG coating had two different purposes during experiments. On one hand, surface-bound long-chain PEG molecules passivated the glass slide by preventing non-specific interaction with proteins and DNA. On the other hand, the sparsely distributed biotinylated-PEG molecules provided a means to immobilize fluorescent probes through biotin-neutravidin-biotin attachment. For the latter purpose, flow chambers were first rinsed with phosphate buffered saline (PBS) and then incubated with a 0.5 mg/ml solution of neutravidin (*Invitrogen*) in PBS for 15 min at room temperature. During this time, neutravidin bound to the biotin molecules exposed by the PEG coating and after flushing

out excess neutravidin, biotinylated fluorescent probes could be attached to the surface bound neutravidin.

2.3. Acquisition, processing and analysis of smFRET data

Data was acquired in the form of movies using the *Andor SOLIS* software (V4.15, *Andor*). Thereby, the electron multiplier gain of the EMCCD camera was enabled and set to 225, the pre-amplifier gain was set to 5x and the readout rate was usually set to 3 MHz at 14 bit. Generally, in smFRET experiments for the NPS analysis of Pol II open complexes (section 3) and of Chd1-nucleosome complexes (section 4.4), short **sif**-movies were recorded with a total duration of 40 s and an integration time of 100 ms per frame (10 Hz frame rate, 400 frames). However, when nucleosome remodeling by Chd1 was investigated (section 4.3), long movies were acquired with a total of 1200-2000 frames to follow the FRET dynamics over an extended period of time. The integration time was set to either 100 ms yielding a total movie duration of 120-200 s, or, in order to observe faster dynamics, to 33 ms resulting in a total movie duration of 40-67 s. Long movies were saved as **tif**-files due to the smaller size compared to **sif**-files.

The acquired data was analyzed using two different custom-written MATLAB (*The Math-Works*) programs. Short **sif**-movies were processed and analyzed with the software *SM FRET V5.7* written primarily by the former PhD student Robert Lewis and a very detailed description of the software can be found in his dissertation [196]. Long **tif**-movies were processed and analyzed with the software *Trace Intensity analysis (TRACY)*, which was developed in the laboratory of Prof. Don Lamb (LMU Munich) by the PhD students Gregor Heiss and Martin Sikor. An elaborate description of TRACY can be found in the PhD thesis of Gregor Heiss [201].

Both programs had a similar workflow with a fully automated routine to find FRET pairs, extract fluorescence intensities of donor and acceptor fluorophores, calculate the local background and finally correct the fluorescence intensities. Subsequently, individual FRET pair fluorescence trajectories could be examined manually and FRET efficiency trajectories were computed for selected traces. In the following, I will describe in more detail the different steps of the workflow, thereby addressing differences between the two analysis programs.

Mapping

In the used TIRF setup, the fluorescence signals of donor and acceptor were spectrally separated with the help of dichroic mirrors and the separated images were projected onto different regions of the EMCCD chip (or later during this work onto two different cameras). Therefore, the first step of data processing was to overlap the images of both channels. Since non linear distortions introduced through optical aberrations were always present, complete overlap could only be achieved by application of an image transformation map

consisting of translation and dilation operations. To create such map, fluorescent beads with a broad emission spectrum, which could therefore be detected in both channels simultaneously, were immobilized in the measurement chamber and imaged. In the resulting images of both channels, fluorescent spots were localized and spots belonging to the same molecule were identified. From matching pairs of fluorescent spots, a transformation map was generated. The identification of corresponding fluorescent peaks was done manually in the *SM FRET* program, but automatically with the option to manually interfere in *TRACY*.

Identification of molecules

In the next step, a peak-finding algorithm was applied to detect the positions of single fluorescent molecules in the movies. Slightly different strategies were applied by the two software solutions. In *SM FRET*, the first 25 frames of the acceptor channel and frames 16-38 of the donor channel were averaged and the absolute minimum of counts per pixel of the entire movie was subtracted from each pixel. The local background was determined by identifying the local minimum in each square of a 16 x 16 pixel grid and this coarse grid of local minima was homogenized through a 20 x 20 pixel image filter and subtracted from the actual image. After scaling of intensities to a range from 0-255, the image was scanned pixel by pixel and all intensities below a certain threshold intensity (manually selected, between 95 % and 99.5 % of the absolutely highest intensity) were set to 0. Then, the image was scanned by a 7 x 7 pixel mask box to identify the location of local maxima with intensities higher than the threshold. Molecules that did not have a round shape or were in proximity to another molecule were discarded. It is important to note that the search for molecules was performed in the red acceptor channel after green excitation. (In ALEX movies, molecules were searched in the red channel after red excitation and in the red channel after green excitation and only if a molecule was present in both, it was considered.) The corresponding molecule in the donor channel was determined after application of the transformation map. In turn, in the *TRACY* software, a threshold was directly applied to the averaged image (average over a user-defined number of frames) of donor and acceptor channel by setting all but the 5 % brightest pixels to zero. Within this threshold image, the pixels with maximum intensities were identified as the center of molecules. This peak search was performed in both, donor and acceptor channel and only those molecules were considered that were found in both channels and could be overlapped using the transformation map. In the case of ALEX movies, acceptor molecules were searched for in the red channel after red excitation (direct excitation frames).

This deviation in the peak-finding algorithm represented the major difference between the two analysis programs. As a consequence, *SM FRET* was found to be better suited to the analysis of samples with a low number of molecules exhibiting FRET and noisy signal. Another difference between both programs was that *SM FRET* analyzed each movie as a whole, whereas *TRACY* divided each movie into segments that were analyzed

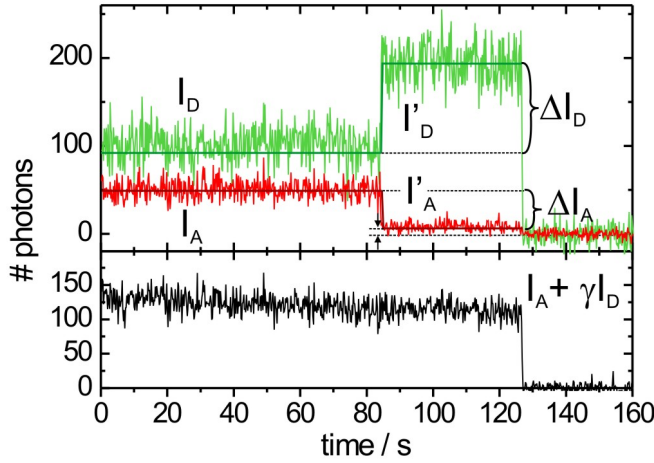


Figure 2.3.1.: Exemplary static FRET time trajectory with all parameters used for γ - and β -factor determination. Fluorescence intensities of donor (green) and acceptor (red) are shown together with the γ -corrected total intensity (black).

sequentially. Therefore, *TRACY* was better suited to the analysis of long movies recorded in the case of bright, long-lived molecules.

Extraction of fluorescence intensity time traces

In the next step, the intensity time traces of donor and acceptor fluorescence were extracted. In *SM FRET*, the intensities within each 7×7 pixel box were summed up for each frame for donor and acceptor and corrected by subtracting the local background. In *TRACY*, the counts within a circular area (radius 3 pixels) around the detected maxima were averaged and background corrected by subtracting the average counts of a ring around the molecule (3 pixels inner radius, 5 pixels outer radius).

Calculation of corrected FRET efficiencies

The FRET efficiency E of each individual FRET pair can be calculated from the background-corrected fluorescence intensities of the donor I_D and the acceptor after donor excitation I_A as

$$E_{\text{FRET}} = \frac{I_A - \beta \cdot I_D}{I_A + \gamma \cdot I_D}. \quad (2.12)$$

The γ -factor accounts for the relative detection efficiencies (η_A and η_D) in the different channels and for differences in the quantum yields (Q_A and Q_D) of acceptor and donor

and is given by

$$\gamma = \frac{\eta_A \cdot Q_A}{\eta_D \cdot Q_D}. \quad (2.13)$$

It can be determined experimentally by comparing the fluorescence intensities before (I) and after (I') acceptor bleaching (Figure 2.3.1) as

$$\gamma = \frac{I_A - I'_A}{I'_D - I_D} = \frac{\Delta I_A}{\Delta I_D}. \quad (2.14)$$

The β -factor accounts for the leakage of donor emission into the acceptor channel and can be expressed as

$$\beta = \frac{I'_A}{I'_D}. \quad (2.15)$$

Both correction factors were determined experimentally for each individual FRET pair. In traces, where the acceptor bleached before the donor, γ and β could be calculated from the respective time-averaged fluorescence intensity signals before and after acceptor bleaching. The corrected total intensity has to be constant until the donor bleaches (Figure 2.3.1, lower panel). The *SM FRET* software only allowed for the analysis of traces, in which the acceptor bleached before the donor. Traces, in which the acceptor did not bleach before the donor, were discarded. In addition, for each molecule a constant γ -factor was assumed and thus, traces that showed dynamics in FRET could not be analyzed since the γ -factor is likely to change when the local environment of the fluorophore alters due to conformational dynamics. However, *TRACY* allowed for the analysis of static as well as dynamic FRET traces. For static traces, a constant γ -factor was assumed and the correction factors were determined as described above. For dynamic traces, the γ -factor was obtained by minimizing the standard deviation of the total intensity to fulfill the following equation

$$\beta = \frac{d}{d\gamma} \text{std}(I_A + \gamma \cdot I_D) = 0. \quad (2.16)$$

Finally, *TRACY* also allowed to include molecules that showed a static signal, but donor bleaching before acceptor bleaching. For such molecules, γ and β cannot be determined, however in this case the average values from other molecules, for which the correction factors could be determined, was used to calculate the FRET efficiencies.

In the calculation of FRET efficiencies, direct excitation of the acceptor by the green laser was neglected, since it was less than 2% of the donor intensity and therefore including a correction factor did not alter the computed FRET efficiencies.

FRET efficiencies were calculated for every time point of a time trace, in the case of static traces after filtering the original data set using a 5- or 10-point sliding average filter. For every molecule, also a mean FRET value was calculated. From the FRET efficiency data,

”framework” FRET efficiency histograms including every time point of several molecules as well as ”moleculewise” FRET efficiency histograms including the average value for each molecule were prepared. Finally, histograms were fit with one or more Gaussians and the maximum FRET efficiency as well as the standard error was extracted from the fit. A more detailed description of the way smFRET data was acquired and analyzed in the individual studies of this work can be found in the experimental procedure sections 3.6 and 4.6).

2.4. Hidden Markov Modeling (HMM)

A *Markov process* is a stochastic, time-dependent process, in which the future of a system with discrete and directly observable states only depends on the present state but not on past states. It is hence without memory and the state-to-state transitions are governed by mono-exponential decay kinetics. Consequently, when a Markov process is modeled to experimental data, the transition probabilities are the only unknown variables. A Markov process becomes ”hidden”, when the individual states are not observable anymore because they are obscured by noise. Therefore, in a *Hidden Markov Model* (HMM), not only the transition probabilities between states but also the states themselves are unknown.

Developed in the 1960s [202], HMM was initially applied to speech and pattern recognition [203, 204, 205]. However, soon after it was widely used for temporal pattern recognition in many fields. It has been applied to bioinformatics questions such as protein and DNA sequence identification [206, 207], pattern recognition in electro-cardiographic curves [208], and biophysical problems such as modeling the current through ion channels [209] and, more recently, analysis of nucleosome unwrapping under force [210]. In 2006, the application of HMM was extended to single molecule fluorescence techniques, where it was used to quantify the number of fluorescently labeled subunits of a polymer by identifying the number of bleaching steps in fluorescence intensity trajectories [211] as well as to monitor RecA filament formation on DNA by smFRET [212, 213].

smFRET time trajectories are usually noisy due to instrumental noise, e.g. shot noise. If dynamic systems are investigated, in which molecules switch between different states due to conformational changes, it is difficult to determine which state is captured by a molecule at a certain time since transitions might be obscured by the noise. Traditionally, states within dynamic smFRET traces were identified either manually ”by eye” or by using a thresholding algorithm, which requires the user to set cutoffs between FRET states. Both approaches suffer from a subjective bias and are impossible in the case of complex systems with more than three states. In contrast, HMM overcomes the shortcomings of these traditional approaches and allows for a user-independent and fully probabilistic analysis of dynamic smFRET traces yielding reliable and reproducible results [212]. In this thesis, I used HMM to extract dynamics from smFRET traces acquired by TIRF microscopy in a similar way as published previously [212]. HMM analysis was performed

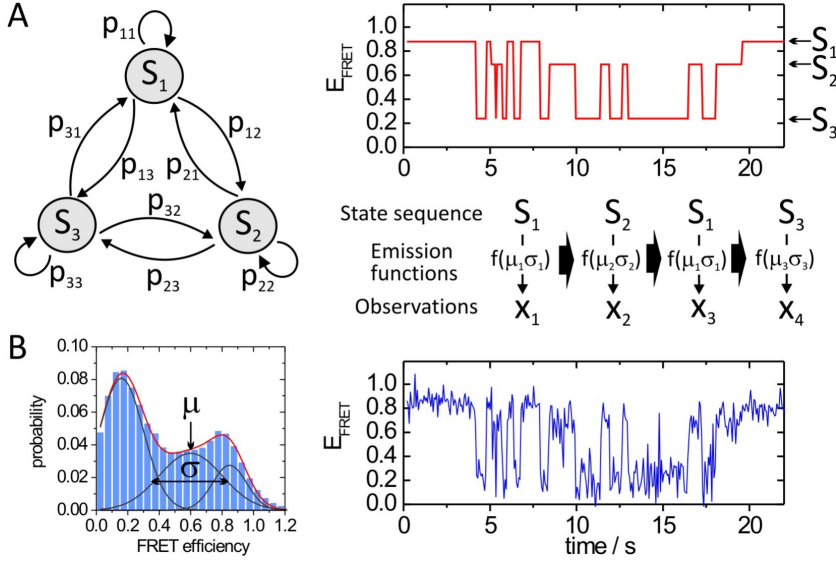


Figure 2.4.1.: Schematic of a single molecule hidden Markov process with three states. (A) A molecule exists in three different FRET states S_1 - S_3 and the probability that it transits at each time point to the other states is given by p_{ij} (i , initial state, j , final state). Each state S_i transits with a probability p_{ii} to itself. Without noise, the system would show an ideal 3 state FRET trace (right, trace identified by HMM analysis from the trace in B). However, in reality a given state sequence $Q = \{q_t\} = S_1, S_2, S_1, S_3$ is hidden by noise and through the emission function $f(\mu, \sigma)$ it generates the observation sequence $X = x_1, x_2, x_3, x_4$. (B) In reality, FRET states show a broad distribution (left, Gaussian distributions with mean value μ and width σ), which masks the idealized sequence of states thereby hiding the underlying Markov process (right). The FRET states μ , the width of distributions σ and the transition probabilities p_{ij} are unknown and have to be determined by HMM. Figure adapted from [201, 212]

using the costum-written MATLAB (*The MathWorks*) software *TRACY* (see section 2.3 and [201]), with an HMM implementation based on the MATLAB HMM toolbox by Kevin Murphy [214].

A molecule that switches dynamically between N discrete conformational states S_i , can be described by the time series $Q = \{q_t\}$ (Figure 2.4.1, 3-state system). For a Markovian chain, the state q_t only depends on the prior state q_{t-1} . The probability for transitions between two states is given by the transition probability matrix P

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \quad (2.17)$$

with $p_{i,j}$ as the transition probability from state i at time t_0 to state j in the next time step $t_0 + \Delta t$. The time step Δt is defined as the integration time of the experiment. The sum of all probabilities of transitions originating from one state has to be 1. The transition probabilities are related to the transition rates by $K = (P - I)/\Delta t$, where I is

2 Physical basics

the identity matrix and K is the rate matrix.

In a hidden Markov process, noise obscures the state sequence $Q = \{q_t\}$ and by the so-called emission function $f(x|\mu_i, \sigma_i)$, a sequence of observations $X = \{x_t\}$ is generated and is detected in the measurement (Figure 2.4.1A bottom, right and B). The emission function describes the distribution of observed emissions from a given state S_i , which is assumed to be Gaussian with a mean value μ and a width σ .

The HMM model M consists of the following parameters:

- The transition probability matrix P .
- The prior π , that contains the prior knowledge about the system, e.g. the distribution of initial FRET values.
- The number of states N and the distributions of output, i.e. N Gaussian distributions with mean μ and width σ .

With a given set of these parameters, the likelihood that the observation $X = \{x_t\}$ can be described by the model M , $P(X|M)$, can be calculated (see [201] for details). This likelihood is the product of all individual step-by-step probabilities $prob(t)$ of observing a molecule in the states $q(t)$, $q(t+1)$ at the time steps t , $(t+1)$, respectively, as given by

$$P(X|M) = \prod_{t=1}^T prob(t), \quad (2.18)$$

with

$$prob(t) = f(x|\mu_i, \sigma_i) \times p(q(t), q(t+1)). \quad (2.19)$$

During the analysis, the HMM is trained, which means that its parameters are modified such that the likelihood $P(X|M)$ is maximized. After many iterations of training, the HMM is optimized in a way that the underlying hidden state sequences Q with the respective parameters (states S_i and transition probabilities p_{ij}) can be retrieved. During the training, only N is kept constant at the value that is estimated by the user beforehand. For the initial states, a flat prior is used meaning that every state is equally probable.

In the used HMM software implemented in *TRACY* (section 2.3), the *Viterbi* algorithm [215] is used to calculate the likelihood $P(X|M)$ for a given set of parameters. The Baum-Welch algorithm [216] is then applied to perform model training and parameter optimization. HMM analysis was performed either globally with one parameter set to describe all trajectories of a given sample or locally with an individual parameter set for each trajectory. If a sample is expected to be heterogeneous in that it consists of different populations of molecules that show distinct dynamics not distinguishable by eye, local HMM analysis should be applied preferentially. However, for all other cases, global analysis should be used.

After HMM analysis, each transition can be plotted as a two-dimensional Gaussian in a

transition density plot (TDP, see Figure 3.3.6C) with the initial FRET efficiency before the transition and the final FRET efficiency after the transition as y and x coordinates, respectively. The standard deviation due to noise found by HMM analysis, σ , is plotted as width of the two-dimensional Gaussian. The volume under the curve of the 2-dimensional Gaussian is normalized to 1 for each transition. The TDP was used in order to extract information about the dwell times. Therefore, a transition population in the TDP was selected and the dwell times of the state before the transition were plotted in a cumulative dwell time histogram (see Figure 3.3.6D) showing the number of transition events with a dwell time longer than a given time as a function of time. The cumulative dwell time distribution could be fitted with a mono-exponential decay (because mono-exponential kinetics are a presumption in a hidden Markov process) to extract the rates of the transitions.

A more detailed description of the settings used for HMM analysis in the individual studies described in this thesis can be found in the experimental procedure sections 3.6 and 4.6).

2.5. Bayesian data analysis and the Nano-Positioning System

smFRET can provide real-time distance information between a single pair of dye molecules within a biological macromolecular complex and can therefore be used for the structural analysis of such complex. If one measures smFRET between a fluorophore attached to an unknown position of the complex (called *antenna* in the following) and three other fluorophores attached to known positions of the complex (called *satellites* in the following), by simple trilateration of the smFRET derived distances the unknown position can be determined. The idea of trilateration by means of smFRET has been applied to numerous biological questions and provided important insights not obtainable by standard structural approaches [15, 16, 17, 18, 19, 217, 218]. However, simple trilateration can only determine one position from the intersection of the individual distances (which is not necessarily the most likely position) and experimental uncertainties affecting this position cannot be taken into account. Possible error sources that contribute to the localization uncertainty are measurement errors of the FRET efficiencies, limited information of the satellite positions due to the attachment of fluorophores to the macromolecule via flexible linkers, and finally the unknown fluorophore orientations. This last uncertainty is most dramatic, since depending on the average orientation of transition dipole moments of donor and acceptor, the Förster distance for a given FRET dye pair may vary considerably, which in turn strongly affects the computed distance (see section 2.1.4 for physical relations between dipole moment orientation, Förster distance and dye-to-dye distance). Another limitation of simple trilateration is that no more than three distance measurements can be included in the localization analysis.

In order to be able to account for these uncertainties and allow for an accurate FRET-based localization of unknown positions within macromolecular complexes, a novel prob-

ability -based analysis method was developed in our laboratory by Adam Muschielok, the so-called Nano-Positioning System (NPS) [20]. In this method, Bayesian parameter estimation [219], smFRET data and x-ray structural information is used to compute a three-dimensional probability density function (PDF) for the unknown antenna position and orientation. In this way, this method does not only allow quantitative FRET-based localization, but also provides an accurate estimate of the localization uncertainty. NPS has been successfully applied in our laboratory to the analysis of yeast RNA polymerase II elongation complexes revealing the position of the 5'-end of exiting RNA [15], the influence of TFIIB on RNA position [20], and the course of nontemplate and upstream DNA [21]. Moreover, it was used to determine the position of transcription initiation factor TFE in archaeal RNA polymerase initiation complexes [220]. Recently, it was first applied outside of our group in the laboratory of T. Ha to reveal structural transitions of a helicase-DNA complex during ATP hydrolysis [221].

While NPS analysis was initially performed for every individual antenna within a macromolecular complex separately, it was recently extended to smFRET networks and global analysis of the data obtained by smFRET (global NPS), which showed an improvement in the localization accuracy [22]. In addition, NPS docking was introduced as a new feature allowing for the docking of macromolecules of known structure constituting a structurally unknown complex.

In this thesis, I used global docking NPS analysis to uncover the three-dimensional architecture of the RNA polymerase II open promoter complex (chapter 3) and to position the chromatin remodeler Chd1 relative to a nucleosome core particle (chapter 4.4). In the following, I will give a brief introduction into Bayesian data analysis and then outline, how it is applied to the problem of smFRET-based localization. Beyond that, an elaborate description of the NPS can be found in the PhD thesis of Adam Muschielok [222].

2.5.1. Probability calculus

Bayesian data analysis is based on the mathematical representation of a state of knowledge or degree of belief as a probability or probability density. This definition of probability is fundamentally different from the common definition of a probability as a relative frequency of an event. Central to Bayesian data analysis is Bayes' theorem [223], which will be derived in the following.

The probability, i.e. the belief that a statement X is true given the background information I can be expressed as $p(X|I)$. Moreover, the probability $p(\bar{X}|I)$ that the statement X is false given the background information I can be directly obtain from the *sum rule*, the first axiom of probability calculus:

$$p(X|I) + p(\bar{X}|I) = 1, \quad (2.20)$$

with $p(\text{false}) = 0$ and $p(\text{true}) = 1$ defining the certainty that something is either false or true.

If Y is another statement, the probability $p(X, Y|I)$ that X and Y are true given the background information I can be calculated as the product of the probability that Y is true and the probability that X is true given Y is true,

$$p(X, Y|I) = p(X|Y, I)p(Y|I). \quad (2.21)$$

This is the second axiom, also called *product rule*. When the expansion is done with X and Y interchanged ($p(X, Y|I) = p(Y, X|I)$), *Bayes' theorem* can be directly derived as

$$p(Y|X, I) = \frac{p(X|Y, I)p(Y|I)}{p(X|I)}. \quad (2.22)$$

2.5.2. Bayesian parameter estimation

To apply the Bayesian approach to data analysis, a physical model has to be defined to describe the performed experiments. This model is defined through its parameters θ and the outcome of the experiments is quantified by data d . Bayes' theorem can now be used to extract information about the parameters in the light of the measured data, if X and Y are replaced by θ and d :

$$p(\theta|d, I) = \frac{p(d|\theta, I)p(\theta|I)}{p(d|I)}. \quad (2.23)$$

Thereby, $p(\theta|d, I)$ is the so-called *posterior*, which describes the probability that a certain set of parameters is true given the data. The probability density $p(d|\theta, I)$ is called *likelihood* and it describes the expectation of how the data would look like if it had been simulated with the given set of model parameters. Moreover, $p(\theta|I)$ is called *prior* and encodes the state of knowledge about the parameters prior to the experiment, i.e. without the knowledge about the data. Finally, the denominator $p(d|I)$ is the evidence, which is constant since the data d does not vary once it is measured. The evidence only becomes important, when different models are compared that could describe the same data.

If several parameters define a model but at the end some parameters are of no interest, one can remove the undesired parameters by integrating over all possible values, a process referred to as *marginalization*. Marginalization is illustrated in the following example: A measured signal s consists of an amplitude a and background b and the background is determined in a reference measurement yielding the data r , then the posterior is given as $p(a, b, |s, r, I)$ and can be calculated by using Bayes' theorem as described above. If the background b is not of importance it can be removed by marginalization:

$$p(a|s, r, I) = \int db p(a, b|s, r, I). \quad (2.24)$$

2.5.3. Bayesian parameter estimation in the global NPS

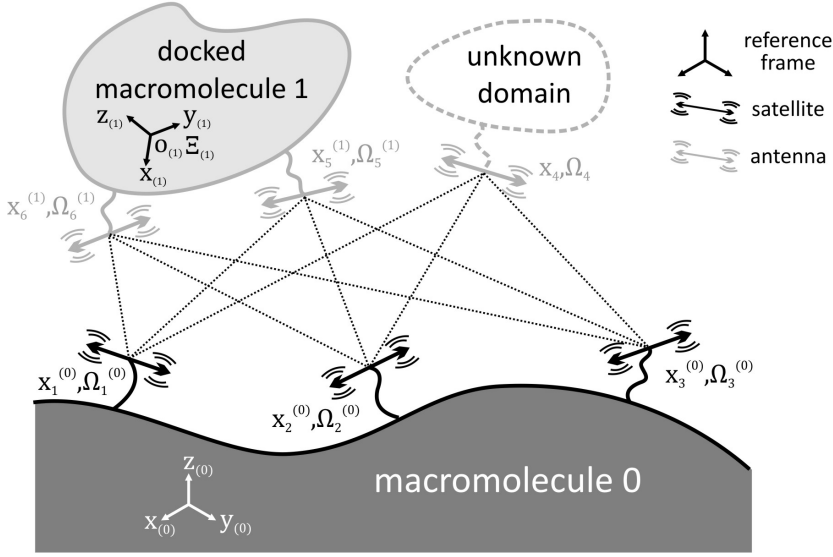


Figure 2.5.1.: Schematic illustration of a FRET network and global docking NPS model parameters. The FRET network is similar to those analyzed in this thesis in chapter 3 and 4.4. The structures of macromolecules 0 and 1 are known, but their relative position and orientation is not and will be determined by NPS docking analysis. The reference frame $(x_{(1)}, y_{(1)}, z_{(1)})$, defined by the origin $o_{(1)}$ and orientation $\Xi_{(1)}$, is assigned to macromolecule 1 and it is positioned and oriented relative to the "root" or "laboratory" coordinate system $(x_{(0)}, y_{(0)}, z_{(0)})$ of macromolecule 0. Three satellites are attached to macromolecule 0 ($i = 1 - 3$) and two antennas to macromolecule 1 ($i = 5, 6$). The fluorophores are anchored to the macromolecules via flexible linkers and are defined by their position \mathbf{x}_i^k and average transition dipole moment orientation $\mathbf{\Omega}_i^k$. The fluctuations around the average orientation are indicated by motion lines. One additional antenna ($i=4$) is attached to a domain of unknown structure and is therefore positioned and oriented independently of a reference frame relative to $(x_{(0)}, y_{(0)}, z_{(0)})$. All measured FRET efficiencies are indicated by dotted black lines. Figure adapted from [222].

Model assumptions and parametrization

To apply Bayesian parameter estimation to smFRET-based localization, a physical model has to be defined describing the experiments. Moreover, model parameters, corresponding priors and a likelihood, which mathematically connects the measured data with the parameters, have to be determined.

Global NPS uses as the model parameter set the positions \mathbf{x}_i and average transition dipole moment orientations $\mathbf{\Omega}_i$ of all fluorophores of a FRET network (see Figure 2.5.1 for a schematic illustration). The observables, i.e. the measured data, are FRET efficiencies $\{E_{ij}\}$, where $\{E_{ij}\} = E_{i1}, E_{i2}, E_{i3}, \dots, E_{iN_{sat}}$ are the FRET efficiencies between the i^{th} antenna ($i = 1, 2, \dots, N_{ant}$) and each of N_{sat} satellites, as well as steady-state fluorescence anisotropies $\{r_{i,\infty}, r_{j,\infty}\}$ of the fluorophores ($i = 1, 2, \dots, N_{ant}$ and $j = 1, 2, \dots, N_{sat}$). FRET anisotropies (or transfer anisotropies) $\{A_{ij}\}$ for each FRET pair can optionally

be included and it was shown to increase the localization accuracy [22]. However, in this thesis only steady-state anisotropies of donor and acceptor are measured and used to calculate the dynamically averaged axial depolarization of the respective fluorophore (see below). Beside the FRET efficiencies and fluorescence anisotropies, the isotropic Förster distances are measured for each FRET pair $\{R_{ij}^{iso}\}$ and are part of the background information I with the assumption that they are known precisely.

The NPS emanates from a static model, which means that the positions of fluorophores \mathbf{x}_i are assumed to be fixed. The crystal structure to which satellite dyes are attached is assumed to be correct. Moreover, the transition dipole moment $\mathbf{\Omega}_i$ of each fluorophore is assumed to perform axially symmetric orientation fluctuations much faster than the time-scale of the fluorescence lifetime. The average transition dipole moment orientation of each fluorophore is assumed to be static and equal for both absorption and emission. The amount of orientation fluctuations is characterized by the dynamically averaged axial depolarization $\langle d_i^x \rangle$ that can be calculated from the steady-state fluorescence anisotropy $r_{i,\infty}$ and the fundamental fluorescence anisotropy $r_{i,0}$ ($r_{i,0} = 2/5$ since the angle between absorption and emission transition dipole moment of a fluorophore is assumed to be 0°) of fluorophore i as

$$\langle d_i^x \rangle^2 = r_{i,\infty} / r_{i,0}. \quad (2.25)$$

The dynamically averaged axial depolarization is used to calculate the dynamically averaged orientation factor $\langle \kappa_{ij}^2 \rangle$ for a given FRET pair consisting of fluorophores i and j as

$$\begin{aligned} \langle \kappa_{ij}^2 \rangle &= (\cos \psi_{ij} - 3 \cos \theta_{ij} \cos \theta_{ji})^2 \langle d_i^x \rangle \langle d_j^x \rangle \\ &\quad + (1/3 + \cos^2 \theta_{ji} \langle d_j^x \rangle) (1 - \langle d_i^x \rangle) \\ &\quad + (1/3 + \cos^2 \theta_{ij} \langle d_i^x \rangle) (1 - \langle d_j^x \rangle), \end{aligned} \quad (2.26)$$

where θ_{ij} , θ_{ji} and ψ_{ij} are the angles that describe the relative orientation of the fluorophores (section 2.1.4). From $\langle \kappa_{ij}^2 \rangle$ the expected Förster distance R_{ij} can be calculated by

$$R_{ij}^6 = R_{ij}^{iso6} \frac{\langle \kappa_{ij}^2 \rangle}{2/3}, \quad (2.27)$$

where R_{ij}^{iso} is the measured isotropic Förster distance for the FRET pair ij for the case that $\langle \kappa_{ij}^2 \rangle = 2/3$. The expected FRET efficiency can then be calculated as

$$\varepsilon_{ij}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{\Omega}_i, \mathbf{\Omega}_j) = \frac{1}{1 + (|\mathbf{x}_i - \mathbf{x}_j| / R_{ij}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{\Omega}_i, \mathbf{\Omega}_j))^6}. \quad (2.28)$$

In the case that docking of macromolecules is performed, the parts of a docked macromolecular complex are regarded as rigid bodies that have a fixed position and orientation

2 Physical basics

relative to each other. In that case, the fluorophore i linked to the k_i^{th} component is uniquely described by its position $\mathbf{x}_i^{(k_i)}$ and average transition dipole orientation $\boldsymbol{\Omega}_i^{(k_i)}$ relative to the reference frame k_i (Figure 2.5.1). With the known reference frame position and orientation, the fluorophore position $\mathbf{x}_i^{(0)}$ and orientation $\boldsymbol{\Omega}_i^{(0)}$ in the root reference frame can be calculated by a simple coordinate transformation.

Likelihood

The likelihood is the product of contributions from each FRET pair, since FRET efficiencies and anisotropies are measured independently. It can be expressed as

$$p(\{E_{ij}\}, \{r_i r_j\} | \{\mathbf{x}_i, \boldsymbol{\Omega}_i\}, I) = \prod_{ij \in M} L_{ij}(\mathbf{x}_i, \boldsymbol{\Omega}_i, \mathbf{x}_j, \boldsymbol{\Omega}_j). \quad (2.29)$$

Here, L_{ij} denotes the contribution of the FRET pair ij as a function of the positions and average transition dipole moment orientations of the respective fluorophores, and M is the set of measured FRET pairs.

Prior

The prior for fluorophore i is given as

$$p(\mathbf{x}_i, \boldsymbol{\Omega}_i | I) = p(\mathbf{x}_i | I) p(\boldsymbol{\Omega}_i | I). \quad (2.30)$$

In the global NPS, structure based geometric constraints are encoded in the position prior. Therefore, a flat position prior $p(\mathbf{x}_i | I)$ is assumed inside the volume accessible to the fluorophores. This accessible volume is simulated using a flexible linker model with a certain linker length, linker diameter and fluorophore diameter for all satellites and antennas attached to a rigid macromolecular reference frame. Thereby, positions inside the known macromolecular structure that cause a sterical clash are excluded. For independent antennas (antenna 4 in Figure 2.5.1), the position prior is defined by a large volume around the crystal structure in the root coordinate system to the exclusion of the van der Waals volume of that structure and it is flat inside this volume. The orientation prior $p(\boldsymbol{\Omega}_i | I)$ is also assumed to be flat and therefore no particular orientation is favored. Details about the position priors used in the individual NPS analyses presented in this thesis can be found in the experimental procedure sections 3.6.8 and 4.6.17.

Posterior

Finally, the complete data set consisting of mean FRET efficiencies, anisotropies, and isotropic Förster distances is used to simultaneously infer positions and orientations of all antennas and the corresponding reference frames within the root coordinate system. This is done by computing the posterior

$$p(\mathbf{x}_i, \mathbf{\Omega}_i | \{E_{ij}\}, \{r_i r_j\}, I) \quad (2.31)$$

with the approximations used for likelihood and prior presented above. Due to the complexity of this inference problem, the computation must be performed fully numerically using a nested sampling algorithm [224] based on Markov chain Monte-Carlo in C and MATLAB (*The MathWorks*). The result is then a set of samples that are used to represent and visualize the posterior. The posterior consists of a probability density function (PDF) for each antenna and each docked reference frame depending on its position and orientation relative to the root coordinate system. In order to obtain the marginal three-dimensional position posterior PDF of the antennas and of any desired position within the docked reference frames, marginalization has to be applied by integrating over all model parameters but \mathbf{x}_i as follows:

$$p(\mathbf{x}_i | \{E_{ij}\}, \{r_i r_j\}, I) = \int d\{\mathbf{x}_{j \neq i}, \mathbf{\Omega}_j\} p(\mathbf{x}_j, \mathbf{\Omega}_j | \{E_{ij}\}, \{r_i r_j\}, I). \quad (2.32)$$

Finally, the marginal position PDFs of the antenna dyes and of any desired position within the docked reference frames relative to the root coordinate system can be exported as `XPLOR`- or `mrc`-files and the credible volumes of the positions can be visualized as iso-surfaces in Chimera [151] or Pymol (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger).

3. Dynamic architecture of a minimal Pol II open promoter complex

3.1. Abstract

In this chapter, I describe the application of smFRET experiments and global NPS analysis to determine the three-dimensional architecture of a minimal Pol II open promoter complex (OC). The studied minimal OC consisted of promoter DNA including a TATA box and an 11-nucleotide mismatched region around the transcription start site (TSS), TATA box-binding protein (TBP), RNA polymerase (Pol) II, and general transcription factor (TF)IIB and TFIIF. Global NPS analysis revealed an architecture, in which TATA-DNA and TBP reside above the Pol II cleft between clamp and protrusion domains. Furthermore, I directly observed the downstream DNA to be dynamically loaded into and unloaded from the Pol II cleft at a timescale of seconds. The TFIIB core domain was displaced from the Pol II wall, where it is located in the closed promoter complex. These results uncovered large overall structural changes during the initiation-elongation transition, which are apparently accommodated by the intrinsic flexibility of TFIIB.

The content of this chapter is the result of a fruitful collaboration with the laboratory of Professor Patrick Cramer and has been published recently [225].

3.2. Introduction

Transcription initiation requires assembly of Pol II with general transcription factors (TFs) at eukaryotic protein-coding gene promoters, melting of DNA surrounding the transcription start site and its insertion into the active center cleft of the polymerase. The resulting open promoter complex (OC) enables RNA synthesis, which triggers release of the general TFs and conversion of the OC to a stable elongation complex (EC). Previous structural studies have elucidated transcription initiation and elongation. The crystal structures of Pol II and the elongation complex (EC) are known [226, 48, 49, 51, 227] and the architecture of the closed promoter complex (CC) was derived by locating general transcription factors on Pol II with the use of biochemical crosslinking [228, 54, 229, 70, 230]. The crystal structure of the Pol II-TFIIB complex could be resolved and led to models of the CC [42, 46]. Models of the open promoter complex (OC) were also proposed based on the Pol II-TFIIB structure and known biochemical data, but the structure of the OC remains unknown. The OC likely contains flexible regions, since it could not be trapped crystallographically, and because conversion from the CC to the EC requires large conformational changes in DNA and general transcription factors.

To structurally analyze flexible Pol II complexes not amenable to x-ray analysis such as the OC, direct visualization in real time is required as provided by single molecule FRET. Here, I used smFRET, global NPS analysis, and NPS docking, to determine the molecular architecture of a minimal OC consisting of Pol II, promoter DNA, TBP, TFIIB, and TFIIF. I measured smFRET efficiencies between antenna dye molecules (antennas) attached to the upstream nontemplate DNA, the TATA box, TBP, and TFIIB, and several 'satellite' dye molecules (satellites) attached to positions on the template DNA and Pol II that are known from crystallographic structures. The data allowed us to build a model of the OC, which provides new insights into the mechanism of transcription initiation.

3.3. Results

3.3.1. Assembly of OCs with labeled DNA and factors

To improve our understanding of the structural intermediates during transcription initiation, I studied a minimal OC preceding elongation and RNA synthesis using sm-FRET measurements and global NPS analysis. I assembled minimal OCs sufficient for promoter-dependent transcription *in vitro* [33] composed of purified endogenous yeast Pol II-TFIIF complex, recombinant TBP, and recombinant TFIIB on a synthetic DNA scaffold with a mismatched region that mimicked the transcription bubble (Experimental Procedures 3.6.3). The scaffold resembled in sequence the one used for x-ray structure determination of the complete Pol II EC [51], but contained a TATA box 30 nt upstream of the TSS and an 11 nt mismatched region around the TSS from register (-9) to (+2). Thereby, +1 marks the TSS and negative and positive numbers denote upstream and downstream positions, respectively (Figure 3.3.1A). In the scaffold, the point of strand separation is located 20 bp downstream of TATA, which corresponds to the distance where DNA opening commences in yeast and human [233, 234]. Application of NPS analysis requires measurement of FRET efficiencies between satellites at known positions within the OC and antennas attached to unknown positions within the OC that are to be determined. Satellites were attached to one of three positions on the downstream template DNA (T-DNA(+3), (+7) or (+12)), or to one position on the upstream template DNA (T-DNA(-10)), or to residue C150 of a single cysteine mutant of the Pol II subunit Rpb7 (Experimental procedures 3.6 and Figure 3.3.1A,C). Antennas were placed on three different key elements of the OC, namely the TBP-TATA DNA complex, upstream DNA, and TFIIB (Figure 3.3.1A,B). TBP-TATA was labeled on one of two positions on the TATA box (NT-DNA(-30) or (-23)) or on residue C128 of TBP (TBP-(C128)). The upstream DNA contained one antenna at position (-18) of the nontemplate DNA (NT-DNA(-18)). TFIIB was labeled with an Sfp phosphopantetheinyl transferase-catalyzed strategy [187] that required insertion of an 11-amino acid ybbR-tag into an unstructured region of the protein. Labels were introduced at two informative positions, one adjacent to the N-terminus of TFIIB cyclin 1 (TFIIB-(122)ybbR), and one in the TFIIB C-terminal tail (TFIIB-(C-term)ybbR) (Experimental procedures 3.6.1).

Introduction of a ybbR-tag into TFIIB did not impair transcriptional activity as shown by an *in vitro* transcription assay carried out by Friederike Hoeg (laboratory of Patrick Cramer) using nuclear extract from a temperature-sensitive yeast strain carrying a point mutation in the TFIIB gene [58] (Figure 3.3.2A; Experimental procedures 3.6.5). Such extracts do not support transcription at 30 °C, but promoter-dependent transcription was recovered when recombinant wild-type TFIIB or one of the two TFIIB-ybbR mutants were supplied. Moreover, I used an electrophoretic mobility shift assay to demonstrate that fluorescently labeled DNA and proteins assembled into stable OCs (Figure 3.3.2B; Experimental procedures 3.6.4).

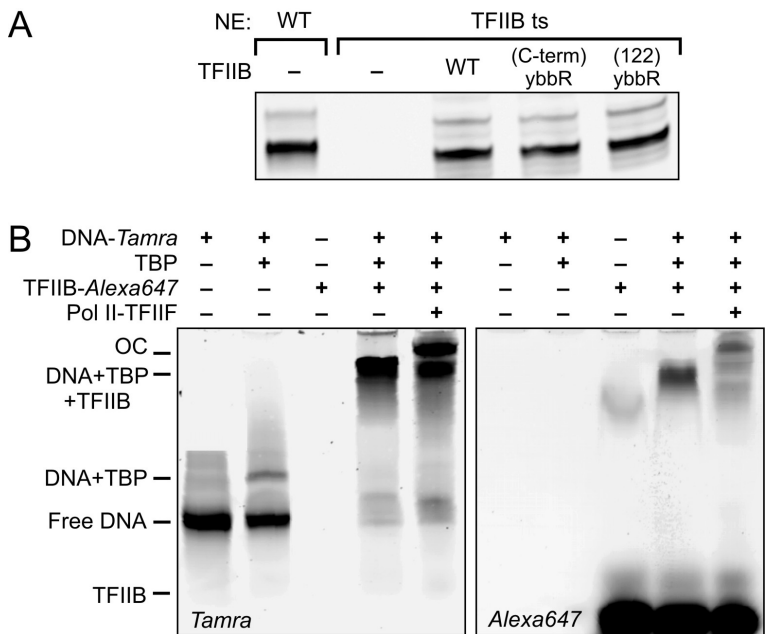


Figure 3.3.2.: *In vitro* transcription of TFIIB-ybbR mutants and Electrophoretic mobility shift assay (EMSA) of labeled OC assembly. (A) *In vitro* transcription assay testing HIS4 promoter activity of TFIIB-ybbR mutants. Analysis of transcription products was performed by primer extension using fluorescently labeled primers (Cy5, excitation: 633 nm, detection: 655-685 nm). Nuclear extract (NE) from wild type (WT) yeast strain showed full *in vitro* transcription activity (lane 1), whereas NE from a TFIIB temperature-sensitive yeast strain was essentially inactive (lane 2). Transcriptional activity could be restored when recombinant WT TFIIB and the mutants TFIIB-(122)ybbR and TFIIB-(C-term)ybbR were supplied (lane 3-5). (B) Fluorescent gel images of OCs prepared from pre-melted DNA bubble, TBP, TFIIB-(C-term)ybbR and Pol II-TFIIF. Left: Tamra labeled DNA was detected in the green channel (excitation: 532 nm, detection: 565-595 nm). To the DNA bubble (lane 1), protein factors were sequentially added: first TBP (lane 2), then Alexa647-labeled TFIIB-(C-term)ybbR (lane 4) and finally Pol II in complex with TFIIF (lane 5). Each additional protein factor produced an additional upshift of the nucleic acid scaffold in the gel showing its binding to the DNA. Right: Alexa647 labeled TFIIB-(C-term)ybbR was detected in the red channel (excitation: 633 nm, detection: 655-685 nm). Lane 3 shows the signal of TFIIB alone. The fluorescent band is shifted up due to the binding of TFIIB to the DNA-TBP assembly (lane 4) and to Pol II in the OC (lane 5). In lane 3-4, residual free Alexa647 from the TFIIB-(C-term)ybbR labeling reaction is visible below the band of free TFIIB. Figure adapted from [225].

3.3.2. Downstream DNA occupies the Pol II cleft

Upon assembly, OCs were purified by size exclusion chromatography, attached to the surface of microfluidic chambers, and investigated by smFRET using total internal reflection fluorescence microscopy (Experimental procedures 3.6.6). The location of the satellites on the downstream template DNA and the template strand in the bubble was assumed to resemble that in the EC [51] (introduced in section 1.1.2), consistent with mutagenesis data [42], and with a recently published crystal structure of Pol II in complex with free DNA [44]. To test this assumption, I measured smFRET between satellites on the template DNA (T-DNA(-10), (+3) or (+7)) as donors and satellite Rpb7-C150 as acceptor for both OC and EC (Figure 3.3.3, Table A.1.1). For the EC, all three measurements resulted in smFRET efficiency histograms showing a single peak. In the OC, the FRET efficiency histogram for the measurement between upstream satellite T-DNA(-10) and Rpb7-C150 resembles the corresponding histogram for the EC, indicating that the position of the upstream template strand in the bubble is not altered. smFRET measurements for the two downstream template DNA satellites T-DNA(+3) and T-DNA(+7) in the OC resulted in bimodal FRET efficiency distributions with two distinct, comparatively narrow peaks that could be fitted with two Gaussian distributions, comprising 70 % and 30 % of the data. The main peak of the distributions resembles the single peak of the EC, indicating that in the major population of OCs the downstream DNA adopts a location in the polymerase cleft similar to that in the EC.

3.3.3. Dynamic loading of downstream DNA into the cleft

In the above described experiments, a minor subpopulation of OCs with a high FRET efficiency was observed, indicating an alternative location of the downstream DNA. Similar observations were made in smFRET measurements from satellites to the antennas positioned on the upstream DNA (NT-DNA(-18), (-23) or (-30), Figure 3.3.4 and Supplemental Figure A.1.1).

For all smFRET measurements including downstream DNA satellites (Figure 3.3.4D-F, Figure A.1.1D-F and K-M) two Gaussians were necessary to describe the data. The side peaks corresponding to a subpopulation of OCs of about 30 % are located at higher FRET efficiencies, indicating that the alternative location of downstream DNA is closer to Rpb7 and upstream DNA. Side populations were also observed in measurements between downstream satellites and antennas on TBP or TFIIB (e.g. Figure A.1.3D, F), however, these data were not included in the following NPS analysis due to comparatively poor statistics. Global NPS analysis (Experimental Procedures 3.6.8) using all side peaks in measurements between positions on the downstream DNA and on the upstream DNA, TATA DNA or Rpb7 revealed that the alternative location of downstream DNA was on top of the cleft between the Pol II clamp domain of Rpb1 and the lobe domain of Rpb2 that flank opposite sides of the cleft (Figure 3.3.5).

smFRET time trajectories showed dynamic transitioning between the two downstream

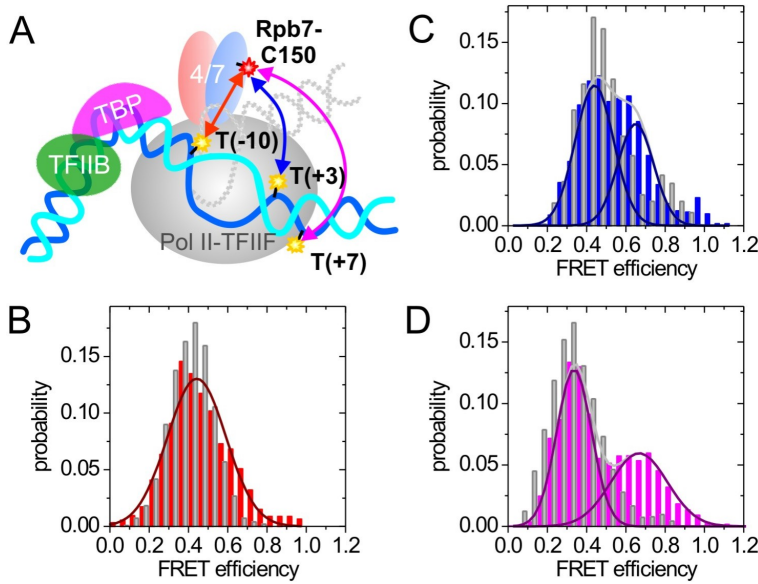


Figure 3.3.3.: Localization of satellites within OC. (A) Cartoon indicating the label positions and smFRET measurements for experiments designed to test, whether the EC structure is a good model for the OC with respect to the position of downstream template DNA and template strand in the bubble. Measurements from satellite Rpb7-C150-Alexa647 to satellites T-DNA(-10)-Tamra (red arrow), T-DNA(+3)-Tamra (blue arrow) or T-DNA(+7)-Tamra (pink arrow) were performed in an OC as well as in an EC. (B-D) Comparison of framewise smFRET histograms from OC measurements (color of histogram matching color of arrow in panel A) and EC measurements (gray histograms). Minor subpopulations are found for the downstream DNA satellites T-DNA(+3) and (+7) at substantially higher FRET efficiencies, which arise from an alternate conformation of the downstream DNA with a shorter distance to Rpb7-C150 (dotted gray lines in panel A). Image adapted from [225].

DNA locations in and above the cleft at a timescale of seconds (Figure 3.3.6). For quantification of the data, I performed Hidden Markov Modeling (HMM) analysis of dynamic FRET trajectories (Experimental procedures 3.6.7). Because of the presence of two peaks in the framewise smFRET histograms, a two-state Hidden Markov process was assumed. Analysis of 88 time traces undergoing 240 transitions resulted in a transition density plot (TDP, Figure 3.3.6C) that shows two distinct transitions from an initial FRET efficiency of around 80 % to a final FRET efficiency of around 30 % and vice versa. From the TDP, cumulative distributions for the dwell times of both transitions were extracted (Figure 3.3.6D) and fitted by mono-exponential decays, which yielded the rates $k_{out} = (0.8 \pm 0.1) \text{ s}^{-1}$ and $k_{in} = (1.5 \pm 0.4) \text{ s}^{-1}$, respectively. The ratio of the rates of movement in and out of the cleft is in good agreement with the respective FRET efficiency histogram (Figure A.1.1K and Table A.1.1), which shows a fraction of approximately 30 % of all OCs in the high FRET state with the downstream DNA above the cleft and a fraction of approximately 70 % in the low FRET state with the downstream DNA localized in the cleft.

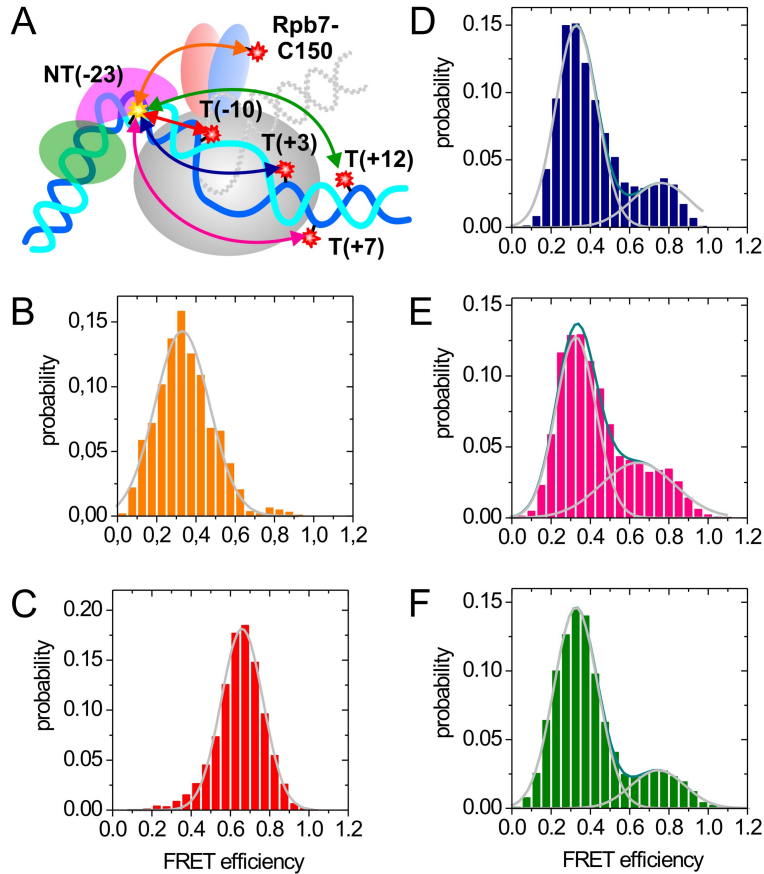


Figure 3.3.4.: Downstream DNA in OC adopts two different conformations. (A) Cartoon illustrating an exemplary set of smFRET measurements from all satellites to an upstream DNA antenna position, namely the data for NPS localization of antenna NT-DNA(-23). The measurements are indicated by double-headed arrows in the same color as the respective FRET efficiency histogram in panel B-F (orange for satellite Rpb7-C150, red for T-DNA(-10), blue for T-DNA(+3), pink for T-DNA(+7), and green for T-DNA(+12)). An OC is presented schematically and the alternate conformation of the downstream DNA is sketched with dotted gray lines. (B,C) smFRET measurements for the satellite positions on the upstream DNA resulted in FRET efficiency histograms, which were fitted with single Gaussians (gray). (D-F) FRET efficiency histograms for measurements to the satellite positions on the downstream DNA were fitted with two Gaussians (gray: individual fit, dark cyan: combined fits). Image adapted from [225].

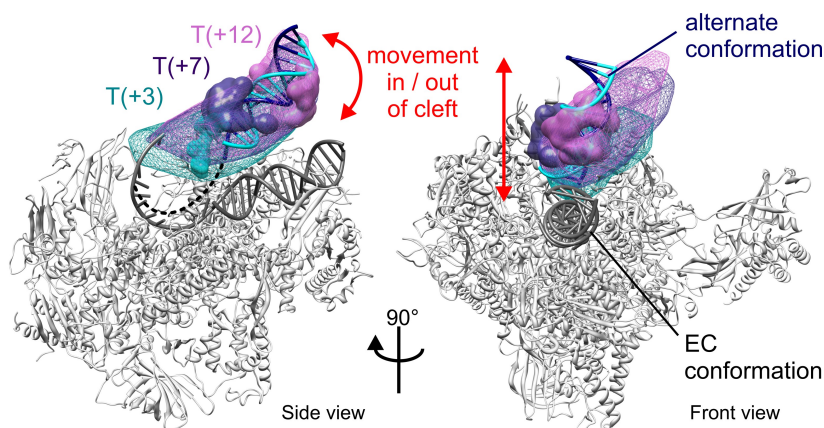


Figure 3.3.5.: Alternate conformation of downstream DNA localized outside of the cleft by global NPS analysis. The position probability densities resulting from the global NPS analysis of the downstream DNA antennas T-DNA(+3) (dark cyan), T-DNA(+7) (dark purple), and T-DNA(+12) (orchid) are shown as meshed credible volumes contoured at 68 % probability relative to the Pol II EC [51]. The conformation of the DNA in the EC (without RNA) is shown in gray. One possible position of the downstream DNA in the alternate conformation (nontemplate DNA, cyan; template DNA, blue) was modeled using Chimera [151]. Thereby, the downstream duplex DNA (register (+3)-(+15)) was moved together with the simulated accessible volumes of the antenna dyes (solid, same color coding as for NPS densities) upwards in the cleft to match the NPS densities. To adopt the alternate conformation, the downstream DNA moves above the cleft between the Pol II clamp and lobe. During this process, part of the single-stranded template DNA upstream of T(+3) is likely to move as well (black dashed line). In the side view, Pol II Rpb9 and a major part of Rpb2 are omitted to better visualize the path of the DNA. Image adapted from [225].

3.3.4. TFIIF stabilizes downstream DNA in Pol II cleft

Since two different conformations of the downstream DNA were observed in the OC but not in the EC, the alternate conformation was supposed to be an effect of the presence of general transcription factors absent in EC measurements. I tested the influence of TFIIF on the downstream DNA conformations by comparing smFRET measurements between all satellites and antenna NT-DNA(-23) on the TATA DNA for OCs with and without TFIIF (Figure 3.3.7). For satellites Rpb7-C150 and T-DNA(-10), unimodal smFRET distributions were obtained that were similar for both measurement conditions and independent of the presence of TFIIF. In contrast, for all downstream DNA satellites a change in the smFRET distributions could be observed regarding the probability ratio of low and high FRET populations, which was dependent on the presence of TFIIF. For OCs lacking TFIIF, the high FRET state corresponding to the alternate conformation of downstream DNA outside the cleft was stronger populated with a probability of ~50 % compared to a probability of ~30 % for OCs containing TFIIF (Table A.1.1). Hence, TFIIF stabilizes the downstream DNA inside the cleft either actively through direct interaction with the DNA or passively via binding to Pol II and thereby modulating the Pol II cleft in a way that downstream DNA is more strongly trapped.

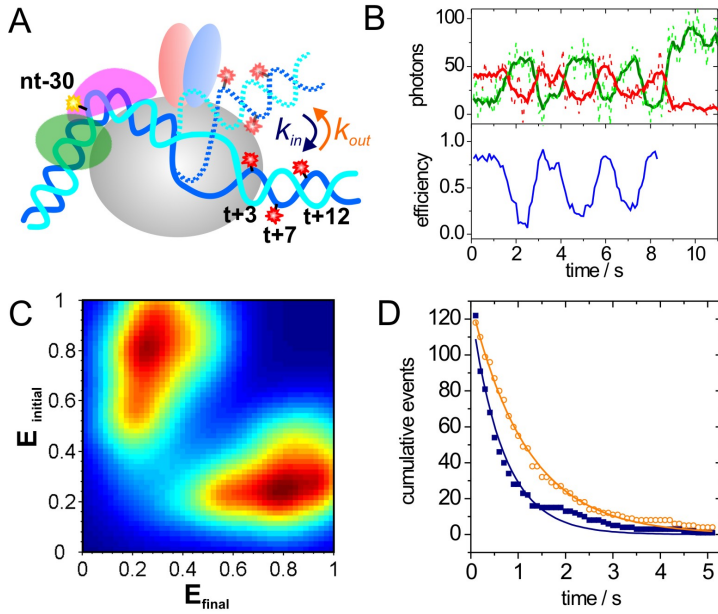


Figure 3.3.6.: Downstream DNA switches dynamically between positions inside and outside of the cleft. (A) Cartoon illustrating the two conformations of the downstream DNA that show dynamic inter-conversion. k_{in} is the rate for the DNA being loaded into the EC conformation (low FRET state) and k_{out} is the rate for the movement out of that conformation into the alternate conformation (high FRET state). (B) Example of directly observed dynamic transitions between low- and high-FRET state. Time trace of fluorescence intensities (dashed lines, raw data, and solid lines, 5 point moving average) for donor NT-DNA(-30)-Tamra (green) and acceptor T-DNA(+3)-Alexa647 (red) are shown together with the computed FRET efficiency (blue). (C) Transition density plot (TDP) for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from HMM analysis of a total of 88 traces showing 240 transitions of the same sample as in panel B. HMM analysis of dynamic traces measured for OCs with NT-DNA(-30)-Tamra and T-DNA(+7)-Alexa647 yielded similar results for the kinetics of the downstream DNA (data not shown). (D) Cumulative distributions showing the number of transition events with a dwell time longer than a given time for the transition from low to high FRET (DNA unloading, orange circles) and for the transition from high to low FRET (DNA loading into cleft, blue squares) using the HMM results shown in panel D. The solid lines represent the corresponding mono-exponential decay fit, from which the kinetic rates $k_{out} = (0.8 \pm 0.1) \text{ s}^{-1}$ and $k_{in} = (1.5 \pm 0.4) \text{ s}^{-1}$ were extracted, respectively. Image adapted from [225].

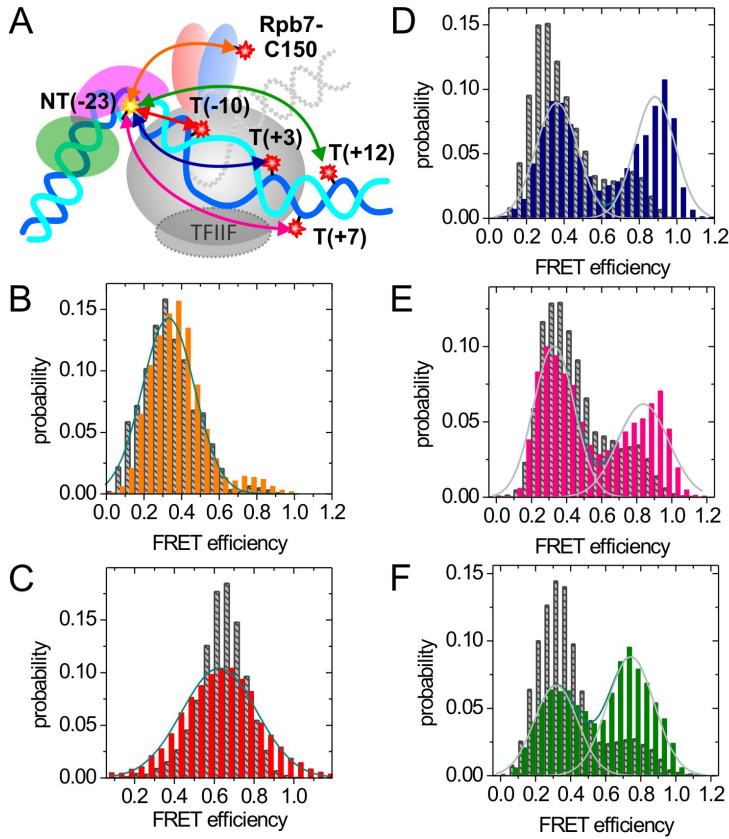


Figure 3.3.7.: TFIIF influences equilibrium between the two different downstream DNA conformations by stabilizing the DNA inside the cleft. (A) Cartoon illustrating the smFRET measurements from all satellites to upstream DNA antenna NT-DNA(-23) performed in the presence and absence of TFIIF (semi-transparent, gray ellipsoid). The measurements are indicated by double-headed arrows in the same color as the respective FRET efficiency histogram in panel B-F. An OC is presented schematically and the alternate conformation of the downstream DNA is sketched with dotted gray lines. (B,C) smFRET measurements for satellites Rpb7-C150 (orange) and T-DNA(-10) (red) resulted in FRET efficiency histograms with a single peak in the presence (gray, shaded histogram) as well as in the absence (colored histogram) of TFIIF. The experiment for T-DNA(-10) in the absence of TFIIF was performed at a time resolution of 10 ms (in contrast to 100 ms in all other experiments) since fast dynamics were observed, presumably as a result of bubble breathing at this position. This explains the broader width of the corresponding histogram compared to the experiment in the presence of TFIIF. (D-F) FRET efficiency histograms for measurements to the satellite positions on the downstream DNA showed two population in the presence (gray, shaded) as well as in the absence (colored blue for T-DNA(+3), pink for T-DNA(+7), and green for T-DNA(+12)) of TFIIF and were therefore fit with two Gaussians (gray: individual fit, dark cyan: combined fits). However, the probability of the high FRET subpopulation of OCs markedly increased in the absence of TFIIF.

3.3.5. TBP and TATA DNA reside above the cleft

To determine the positions of upstream DNA including the TATA box, TBP, and TFIIB within the OC, I measured smFRET for all combinations of satellite and antenna pairs and for several combinations of two satellites or two antennas yielding a total of 37 measurements (Figure 3.3.8, Table A.1.1). I also measured steady-state fluorescence anisotropies of the donor and acceptor dyes for all attachment sites and experimentally determined the isotropic Förster distance for each acceptor and donor dye pair (Table A.1.1). To apply NPS docking [22], two rigid macromolecular structures were defined as reference frames that were to be localized and oriented relative to Pol II (Experimental procedures 3.6.8, Figure 3.3.1). The structures were assumed not to be subject to conformational change when incorporated in an OC. The TATA-TBP reference frame contained antennas NT-DNA(-23), NT-DNA(-30) and TBP-C128, whereas the upstream DNA reference frame contained antenna NT-DNA(-18). Since the structure of TFIIB in the OC is unknown, the antennas on TFIIB were treated as independent dye molecules without the assignment to a reference frame. The position priors, i.e. the accessible volumes, of satellites and of antennas assigned to a reference frame were modeled in the respective coordinate system using a flexible linker model (Experimental procedures 3.6.8). For all antennas without assignment to a docked reference frame, the accessible volume was restricted to a volume not occupied by Pol II in the EC structure, however this Pol II van-der-Waals volume was not excluded in the position prior of docked reference frames.

In the global NPS docking analysis, the complete data set consisting of mean FRET efficiencies, anisotropies, and isotropic Förster distances was used to simultaneously infer positions and orientations of all antennas and the corresponding reference frames within the Pol II coordinate system. As a result of Bayesian parameter estimation, we obtained the three-dimensional probability density of each antenna and of any desired position within the docked reference frames. Figure 3.3.9 displays the densities of the antenna attachment points as 68 % credible volumes in the Pol II coordinate system. Comparison with the CC model derived from crosslinking data [228, 229, 235] and modeling based on x-ray crystallography [42] reveals significant structural differences. Whereas TATA DNA and TBP are located over the wall at the end of the cleft in the CC, the densities for TATA box positions are localized above the cleft between the coiled-coil motive of the Rpb1 clamp and Rpb2 protrusion in the OC. The density of NT-DNA(-18) indicates a shift of upstream DNA into the upper cleft towards the rudder motive of the clamp, and the density for the antenna on TBP is localized above the clamp coiled-coil.

3.3.6. Upstream DNA is stabilized by initiation factors

To investigate whether the observed location of the TBP-TATA DNA complex and upstream DNA is a result of the presence of general transcription factors, I compared smFRET data from antenna NT-DNA(-30) on the TATA box to satellite T-DNA(+7) ob-

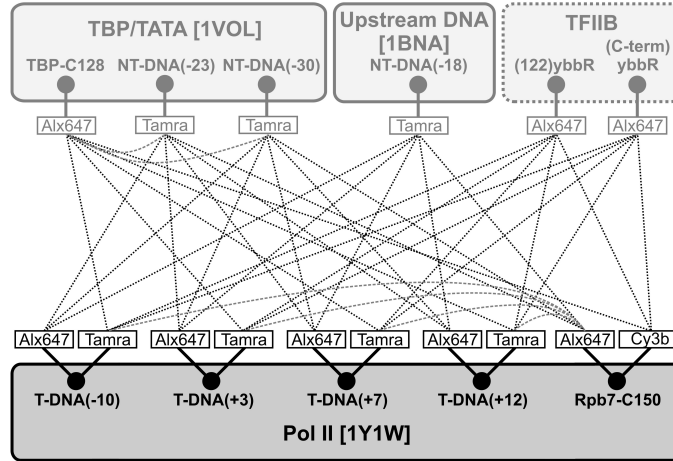


Figure 3.3.8.: FRET network for NPS docking analysis of the OC. The schematic illustration shows all satellites and antennas in their respective reference frame together with the attached dye molecules. FRET efficiencies ($n = 37$) were measured between pairs of satellite and antenna dyes (dotted black lines) and in between satellites or antennas, respectively (dashed gray lines). The location and orientation of TBP-TATA and of antennas NT-DNA(-18), TFIIB-(122)ybbR and (C-term)ybbR (gray-hued) relative to the Pol II coordinate system (dark-hued) were to be determined by global NPS analysis [22]. Image adapted from [225].

tained in an OC, a Pol II-DNA scaffold complex lacking factors, and an EC without factors but comprising a 17 nt RNA (Figure 3.3.10). The recorded smFRET efficiency histograms for the EC and the Pol II-DNA complex were very different from the bimodal OC histogram. Both histograms covered a wide range of FRET efficiencies from 10 to 80 % with the maximum at around 15 % FRET. The examined OC is thus distinct from the EC regarding the location of upstream DNA and initiation factors stabilize the new upstream DNA location. The presence of a side peak in the OC histogram is consistent with the presence of the above mentioned side peaks in measurements between downstream DNA and Rpb7-C150 and results from dynamic loading and unloading of downstream DNA. These results show that the relocation of the upstream promoter assembly from the top of the wall to a defined position above the cleft is a consequence of the presence of the initiation factors in the OC.

3.3.7. Model of the OC

To build a model of the OC structure, we started from the CC model [42] and used the probability densities of the antenna dye attachment points on the TATA box, TBP, and the upstream DNA, together with structural considerations (Figure 3.3.11A). Promoter DNA upstream of register (-10) including TATA box and TBP was moved as a rigid body (arrows in Figure 3.3.9) such that the antenna attachment points on TBP, the TATA DNA, and NT-DNA(-18) were situated in their respective densities, and such that the

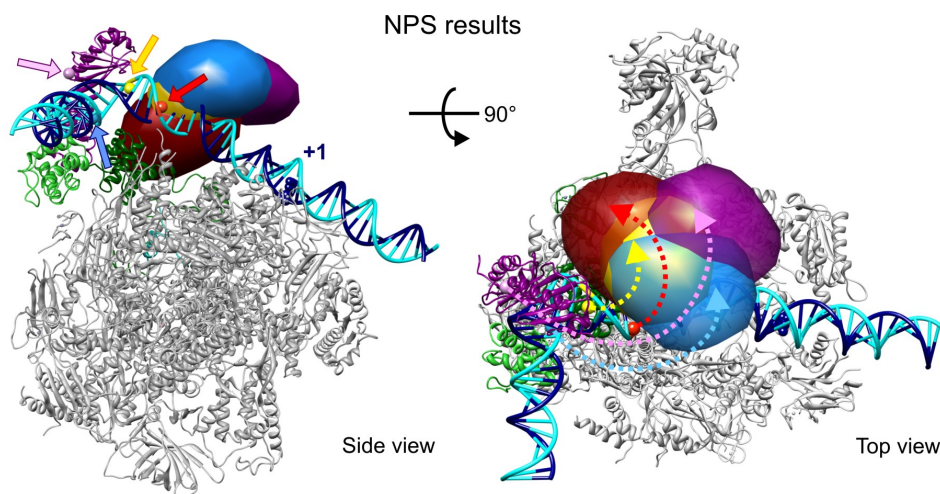


Figure 3.3.9.: NPS results for upstream promoter DNA and TBP. The densities of the antenna dye attachment points at NT-DNA(-18) (dark red), NT-DNA(-23) (yellow), NT-DNA(-30) (light blue), and TBP-C128 (dark magenta) within the OC as revealed by NPS are shown as credible volumes contoured at 68 % probability. The CC model [42] is also shown (NT-DNA, cyan; T-DNA, blue; TBP, dark magenta; and TFIIB, green) with the TSS represented as a space-filling model. For comparison, the antenna dye attachment points on the nontemplate upstream DNA and on TBP within the CC model are shown as spheres (marked by arrows in side view) using the same color coding as for the densities. In the top view, arrows indicate the movement of the different antenna attachment points on DNA and TBP during the transition from CC to OC. Images were prepared using Chimera [151]. Figure adapted from [225].

connection of downstream DNA with the upstream template strand at (-10), the last ordered nucleotide in the Pol II EC structure, was maintained. This resulted in a clash of the upstream DNA base pairs (-12)-(-16) with the B-linker. We assumed that the B-linker was released from the clamp coiled-coil in the OC, and removed it from the model. To avoid minor clashes of the nontemplate DNA registers (-11) and (-10) with the Pol II rudder and fork loop 1, we extended the melted region by two bases upstream. In the resulting model, re-annealing of upstream DNA strands occurs at register (-12), and the backbone of nontemplate DNA at this register resides close to Rpb1 residues 313-316 in the rudder and the conserved Lys471 in Rpb2 fork loop 1. The upstream DNA duplex exits from the Pol II cleft between clamp and protrusion, generally as in the complete EC [21].

3.3.8. B-core is displaced from the wall

Since the position of both, TATA box and TBP, change during the transition from CC to OC, it was important to determine whether the position of the B-core, which mediates the attachment of TATA/TBP to the Pol II wall in the CC, changes as well. Figure 3.3.12 shows the NPS results for the antennas on TFIIB within the Pol II coordinate system

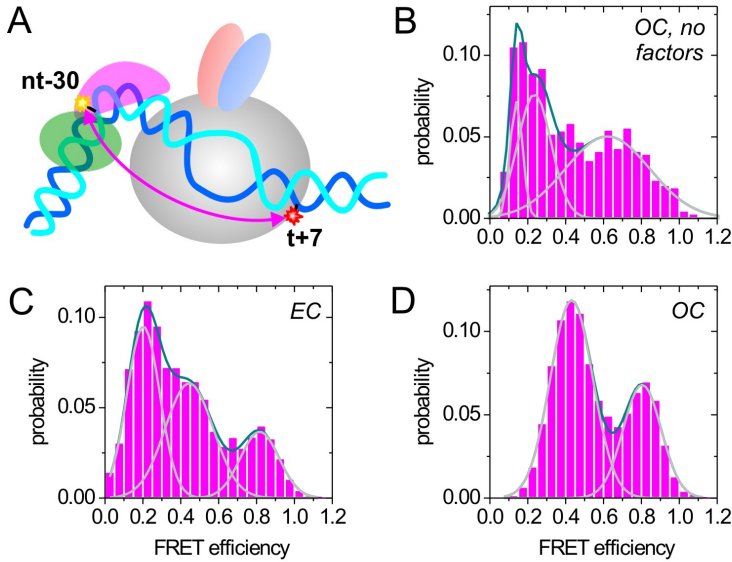


Figure 3.3.10.: Position of upstream DNA in the OC is stabilized by initiation factors. (A) Cartoon illustrating the control smFRET measurements between satellite T-DNA(+7)-Alexa647 and antenna NT-DNA(-30)-Tamra within the EC, the OC lacking all initiation factors and the OC (shown here). (B) Framework FRET efficiency histogram for EC composed of a DNA-RNA bubble with a 17 nt RNA transcript and Pol II. The data covered a broad range of FRET efficiencies and had to be fitted using at least 3 Gaussians (gray, individual Gaussian fits and cyan, combined fit). (C) Framework FRET efficiency histogram for a factor-less OC composed only of DNA bubble and Pol II. Similarly to the data for the EC, the factorless OC yielded a broad FRET efficiency histogram that is best described with three Gaussians. (D) Framework FRET efficiency histogram for the OC. The FRET histogram is different than the histograms for EC and OC lacking all initiation factors showing two distinct FRET efficiency distributions. The side peak is caused by the conformational dynamics of the downstream DNA in the OC. A second distance was investigated in the same way as presented here and yielded similar results (data not shown). Figure adapted from [225].

together with the model of upstream promoter DNA and TBP in the OC. The density for antenna TFIIB-(C-term)ybbR is located above the cleft adjacent to TATA box and TBP on the side facing the downstream DNA. The antenna at the N-terminus of cyclin 1, TFIIB-(122)ybbR, is situated on the same DNA side, but shifted towards the downstream DNA and the Pol II clamp. The NPS densities show no overlap with the modeled accessible volumes of the antennas in the CC model (Figure 3.3.12) and therefore are inconsistent with the position of the B-core in the CC as revealed by crosslinking [228] and x-ray crystallography [42]. Hence, the B-core is displaced from the wall in the OC, but likely still interacts with DNA and TBP. In order to rule out that the difference in the B-core position arises from differences in measurement conditions of smFRET compared to crosslinking or x-ray crystallographic experiments, we measured smFRET between Rpb7-C150 and TFIIB-(122)ybbR within Pol II-TFIIB complexes (Supplemental

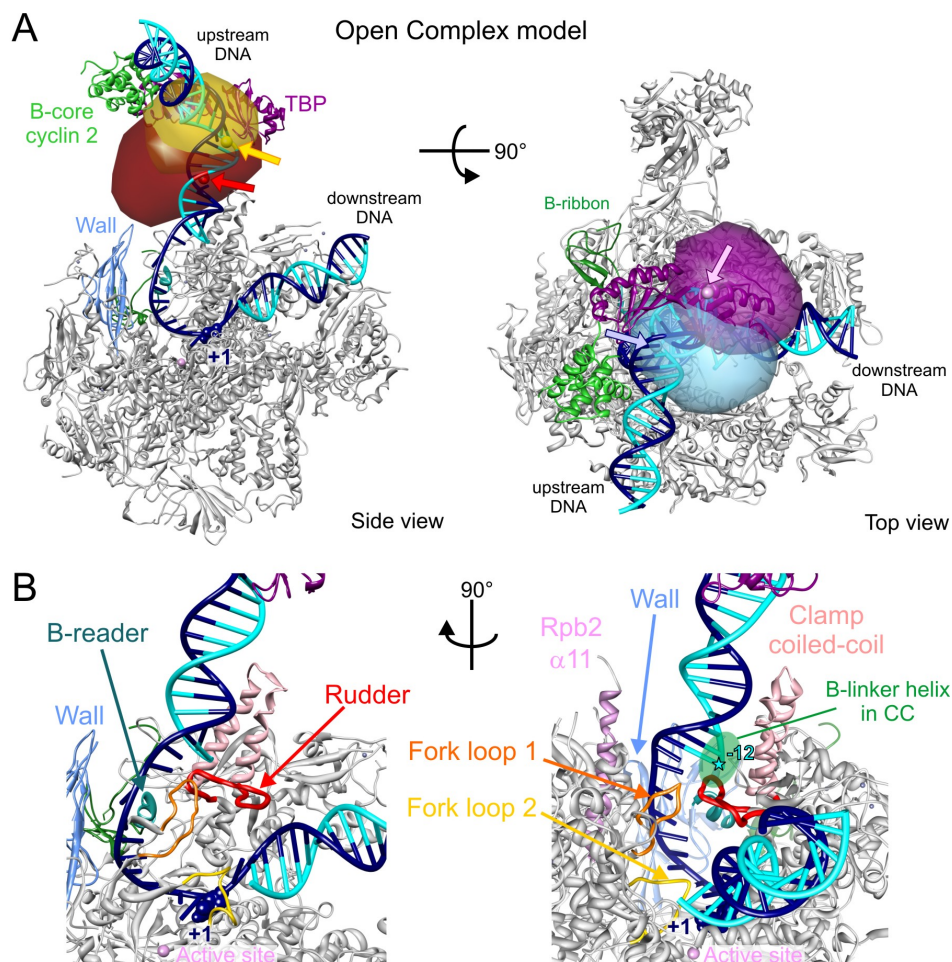


Figure 3.3.11.: Model of the OC based on the results of global NPS analysis. (A) The antenna dye attachment points of NT(-18) and NT(-23) (side view) and of NT(-30) and TBP-C128 (top) are shown as spheres together with the corresponding NPS densities as semitransparent credible volumes (NT(-18), NT(-23): 80 % probability; NT(-30), TBP-C128: 68 % probability). (B) Pathway of DNA in the OC model entering the active center cleft. Zoom into the OC model shows structural features important for OC formation. The TSS is shown as a space-filling model. The position of the B-linker helix in the CC is indicated. All images were prepared using Chimera [151]. Figure adapted from [225].

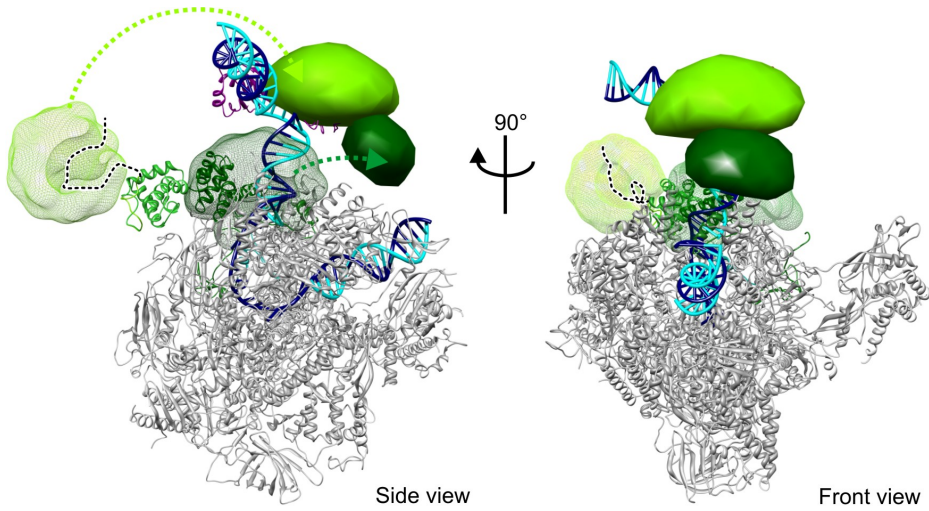


Figure 3.3.12.: NPS localization results of TFIIB. Three-dimensional probability densities of the antenna dyes TFIIB-(122)ybbR (dark green) and TFIIB-(C-term)ybbR (light green) revealed by NPS (68 % credible volumes) are shown in the Pol II coordinate system together with the OC model of the upstream DNA and TBP and with the B-core cyclin folds in the CC model. The meshed volumes represent the corresponding accessible volumes of the antenna dyes within the CC model (same color coding). Because the conformation of neither ybbR-tag nor flexible linker connecting the dye with the ybbR-tag is known, antenna dyes are able to access large volumes. For illustration reasons, I modeled the unknown parts of TFIIB in one possible conformation using coot [232] and indicated them by a dashed line. Arrows indicate the rearrangement, which both antennas need to undergo to match the NPS results. Images were prepared using Chimera [151] and are adapted from [225].

Figure A.1.4). The resulting FRET efficiency and computed distance was consistent with the corresponding distance in the Pol II-TFIIB crystal structure, but different from that in the OC. This indicates that the Pol II-TFIIB complex captures the same structure under both measurement conditions and that the relative position of TFIIB is different in the OC.

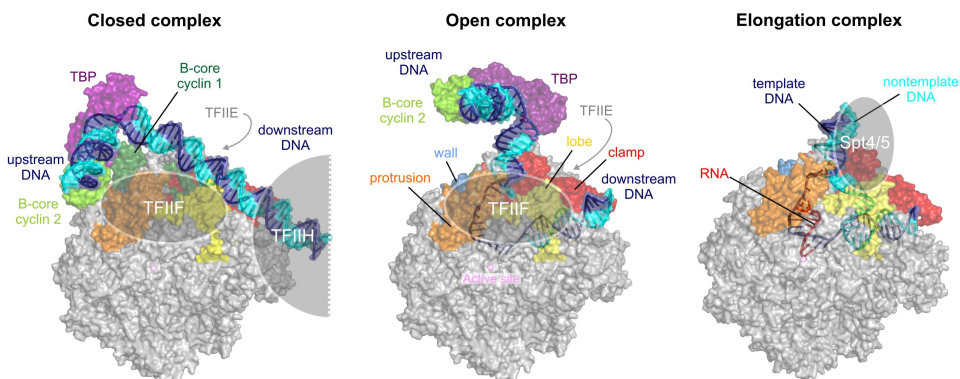


Figure 3.4.1.: OC architecture in the context of the initiation-elongation transition. CC model [42] (left), OC model [225] (middle), and complete EC model [21, 51] (right) are shown as semitransparent surface representations in the side view. Nucleic acids are additionally presented using cartoon representations. Locations of the general transcription factors TFIIIF, -IIH, and -IIE as determined by biochemical probing [229, 70, 230] as well as elongation factor Spt4/5 [237] are indicated schematically. All images were prepared using Pymol (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger). Image adapted from [225].

3.4. Discussion

In this work, the three-dimensional architecture of the dynamic Pol II minimal open promoter complex, a key intermediate in transcription initiation, was determined by sm-FRET experiments, global NPS analysis and modeling based on x-ray crystallographic information. In the resulting OC model, upstream promoter DNA and TBP are located above the cleft, whereas they are located over the wall at the end of the cleft in the CC model (Figure 3.4.1). During the closed-open transition, the TATA-TBP assembly is apparently released from the Pol II wall, in contrast to a prior assumption [42, 46]. Such a rearrangement explains the difference in electrophoretic mobility *in vitro* of the CC and OC [236, 234].

While the positions of upstream DNA and TBP are defined in the minimal OC, the position of the downstream DNA is switching between the location in the cleft that is observed also in the EC, and an intermediary state above the cleft. TFIIIF has a stabilizing effect on the downstream DNA inside the cleft and hence shifts the equilibrium between both conformations towards the EC conformation. The kinetic rates for this DNA loading/unloading process are on the order of a second, thereby presenting a major kinetic trap in this complex assembly process. Similar rates for DNA opening and closing were recently reported for the mitochondrial RNA polymerase of yeast Rpo41 and its transcription factor Mtf1 [12]. The rate of DNA loading into the cleft might be biased by the mismatch present in the DNA scaffold and is likely to change when a fully complementary DNA is used. However, the observation that melted DNA is leaving the cleft cannot be influenced by the mismatches, since at this point DNA is by definition melted. Moreover, transcription factors not included in the minimal OC such as TFIIIE, -H, -D or mediator

are likely to influence the observed dynamics of the downstream DNA, thereby regulating the stability and activity of the OC. I assume that additional factors shift the equilibrium between different intermediate states during the transition from closed to elongation complex by stabilizing or destabilizing successive states. It is important to note that the observed dynamics of downstream DNA are conceivable even in the presence of additional initiation factors.

The results presented here have implications for the mechanism of DNA melting. During DNA melting, torsional stress on the DNA duplex is released [238], and this may involve a counter-clockwise rotation of the upstream DNA-TBP assembly around the TBP C2-axis by about 45° that is suggested by our modeling. *In vivo*, DNA melting around the transcription start site is assisted by TFIIF, which has helicase activity and requires the energy of ATP hydrolysis to open the two DNA strands. Crosslinking studies revealed that TFIIF interacts with DNA exclusively downstream of the TSS, apparently acting as a molecular wrench that rotates downstream DNA relative to a fixed upstream promoter assembly, thereby generating torque and melting the intervening DNA [230] (Figure 3.4.1). Our results suggest that this torque exerted by TFIIF not only melts the DNA but also leads to a release of the upstream promoter assembly from the Pol II wall.

Moreover, the results converge with published data on a better understanding of the OC-EC transition (Figure 3.4.1). First, the course of upstream duplex DNA emanating from Pol II is similar in the OC and EC [21], requiring hardly any change in upstream DNA location during the OC-EC transition. Whereas upstream DNA is fixed in the OC, presumably due to a stabilization by initiation factors, it is mobile in the EC lacking elongation factors [21]. Second, the close proximity of the junction between upstream DNA and the transcription bubble to the rudder (Figure 3.3.11B) is consistent with the essential function of the rudder during OC formation by archaeal and bacterial RNA polymerases *in vitro* [239, 240]. Third, the upstream edge of the transcription bubble may be maintained by TFIIF, which was shown to be positioned on the clamp [229, 220]. At last, during the OC-EC transition, TFIIF is likely displaced by Spt4/5, which was recently shown to bind the clamp coiled-coil in the EC thereby interacting with and stabilizing the upstream DNA [220, 241, 237]. The presence of initiation factor TFIIF in the OC and of Spt4/5 in the EC might induce a slight shift of the path of upstream DNA shown here.

Finally, our results suggest that B-core and B-linker may be displaced from the Pol II surface in the OC, indicating a stepwise release of TFIIB during the initiation-elongation transition. TFIIB remains associated with the complex, apparently via interactions of its B-ribbon and B-reader with Pol II, and aided by stabilizing interactions with TFIIF [242]. Although the many flexible connections within TFIIB prevented us from modeling its structure within the OC, we could show that the B-core resides above the cleft, rather than bound to the wall as in the CC. Our NPS data agree with a model, in which the cyclin 2 domain of the B-core remains bound to the TATA-TBP complex as in the B-core-TATA-TBP crystal structures [243, 55, 244] (Figure 3.3.9A and 3.4.1). In contrast, our

3 Dynamic architecture of a minimal Pol II open promoter complex

data and modeling show that the location of the B-core cyclin 1 domain in the OC is not corresponding to that in the crystallized B-core-TATA-TBP complex. This indicates that the two cyclin domains can move with respect to each other, consistent with NMR data [245]. These results underline the key role of TFIIB and its intrinsic flexibility during the initiation-elongation transition, and explain the experimental difficulties in structurally defining the initiation-elongation transition. Our results are consistent with a recent biochemical crosslinking study showing that the B-core is detached from the wall in minimal yeast OCs, whereas B-ribbon and -reader remain bound to Pol II like in the CC [71]. However, discrepancies exist regarding the role of TFIIF: Fishburn et al. suggest that TFIIB is released from the wall only in complexes lacking TFIIF but remains bound to the Pol II wall when TFIIF is present. In contrast, our results suggest that TFIIB is released from the wall independent of TFIIF. Hence, future studies should aim for elucidating the role of TFIIF during transcription initiation and its putative interaction with TFIIB in order to clarify the existing discrepancies.

3.5. Summary and Outlook

The application of smFRET, global docking NPS analysis, and modeling based on x-ray crystallographic information enabled us to determine the molecular architecture of a dynamic, minimal OC consisting of Pol II, promoter DNA, TBP, TFIIB, and TFIIF. The results uncovered large overall structural changes during the initiation-elongation transition, which are apparently accommodated by the intrinsic flexibility of TFIIB. Moreover, dynamic loading of the downstream DNA into the cleft and unloading from the cleft could be directly observed on a timescale of seconds. The study presents a significant progress towards understanding the mechanism of Pol II transcription initiation, however many open questions remain.

The study presented in this chapter directly suggests further experiments investigating the role of initiation factors in the loading dynamics of melted promoter DNA. The observed loading/unloading dynamics should be used as a starting point to explore how transcription factors such as TFIIF and TFIIE stabilize or destabilize the promoter DNA in the cleft, and which structural domains of the protein factors are responsible.

Regarding TFIIF, a first step has been already taken. Endogenous *S. cerevisiae* TFIIF purified from yeast in complex with Pol II was shown to stabilize the downstream DNA in the Pol II cleft. I verified that recombinant TFIIF is able to reproduce the stabilizing effect on the downstream DNA (Figure 3.5.1). This paves the way for mapping out the regions of TFIIF responsible for the stabilization of promoter DNA, since truncated and mutated constructs of TFIIF can now be tested in a similar way.

Furthermore, the effect of TFIIE and -F on the overall architecture of the OC should be explored and their location within the OC should be determined using smFRET and global NPS analysis.

The nucleic acid scaffold used throughout the experiments presented here consisted of an artificial promoter sequence designed under consideration of the characteristics of a typical mammalian core promoter [246]. However, the sequence of the template DNA strand within the melted region was proposed to have an effect on the architecture and activity of transcription initiation complexes possibly through recognition by TFIIF [71]. Therefore, the architecture of Pol II OCs should be tested using nucleic acid constructs with different, ideally native sequences. Moreover, mammalian and yeast promoters deviate with respect to the distance between TATA box and TSS. Since yeast Pol II and TFs were used in all experiments, the architecture of OCs should be examined using native yeast promoters and compared to the results obtained with the artificial mammalian promoter used here.

Finally, beyond further studies of the Pol II OC, future experiments should use smFRET experiments and global NPS analysis also to explore the architecture of the Pol II closed promoter complex to either verify or correct existing models. Moreover, in the presence of TFIIH, smFRET could be used to follow DNA melting and directly observe the conforma-

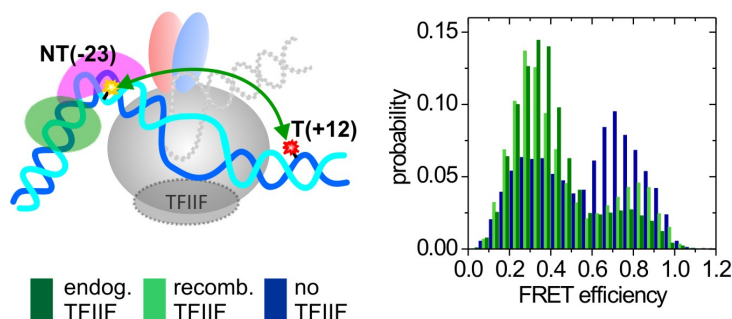


Figure 3.5.1.: Recombinant TFIIIF from *S. mikatae* (tfg1) and *S. cerevisiae* (tfg2) can exhibit similar stabilizing effect on downstream DNA as endogenous TFIIIF. Left: The cartoon illustrates the smFRET measurements from satellite T-DNA(+12) to antenna NT-DNA(-23), which was performed in the presence and absence of endogenous and recombinant TFIIIF (semi-transparent, gray ellipsoid). An OC is presented schematically and the alternate conformation of the downstream DNA is sketched with dotted gray lines. Right: FRET efficiency histograms for the measurements in the presence of endogenous (dark green) or recombinant (light green) TFIIIF as well as in the absence of TFIIIF (blue). Without TFIIIF, the majority of OCs contain the downstream DNA in its alternate conformation outside the cleft (high FRET), whereas in the presence of TFIIIF (both endogenous and recombinant) the majority of OCs contains downstream DNA inside the cleft (low FRET).

tional changes that occur during the transition from closed to open complex. In this way, the kinetics of this dynamic process could be determined. Since TFIIH is a 10-subunit protein factor, such experiments are complex and challenging. As a promising alternative, the closed to open transition can be studied using the archaeal RNA polymerase initiation complex as model system, since it does not require any transcription factor for DNA melting. The archaeal RNA polymerase is both functionally and structurally very similar to eukaryotic Pol II, but has the advantages that it can be prepared as a fully functional recombinant enzyme and that it only requires three transcription factors (TBP, TFB and TFE). As a consequence of the reconstitution from recombinantly expressed subunits, the archaeal RNA polymerase can be site-specifically labeled at any position of the individual subunits and, together with its reduced complexity, it therefore presents a perfect model system for NPS analysis.

Currently, the structural study of Pol II CC using NPS analysis is limited due to a lack of satellite attachment sites within Pol II itself. The position of promoter DNA in the CC is not known, and hence satellites on Pol II are required to localize DNA as well as general TFs. In this study, Rpb4 and 7, which can be recombinantly expressed and reassembled with the Pol II core, were fluorescently labeled via cysteine chemistry. A protocol was recently developed that allows for recombinant expression of Rpb9 and endogenous expression of Pol II(Δ Rpb9) and their subsequent reconstitution [247]. This protocol should be used in future studies to introduce fluorescent labels into Rpb9 of Pol II. Except for Rpb4, 7 and 9, Pol II cannot be reconstituted from its individual, recombi-

nantly expressed subunits *in vitro*, which impedes labeling via single cysteine mutations. Therefore new labeling strategies will need to be exploited based on the introduction of unnatural amino acids or small peptide tags into specific sites of the Pol II structure, in order to be equipped with a sufficient set of satellites on Pol II.

I believe that smFRET experiments and global NPS analysis applied to the analysis of structure and dynamics of Pol II transcription intermediates will continue to contribute substantially to the progress in our understanding of the mechanisms underlying Pol II transcription.

3.6. Experimental Procedures

3.6.1. Protein expression, purification and labeling

TBP

Saccharomyces cerevisiae TBP core with an N-terminal His-tag (amino acids 61-240, gift from Z. S. Juo) and mutant, full-length TBP lacking all native cysteines (C78A, C164A) but holding a single cysteine mutation at residue 128 (S128C) were expressed for 15 h at 20 °C in *E. coli* BL21-Codon plus (DE3) RIL (*Stratagene*). For mutant TBP-S128C, a plasmid for TBP-C61 (gift from M. Meisterernst) lacking all native cysteines (C78A, C164A) and holding a single cysteine mutation at residue 61 (S61C) was subjected to site-directed mutagenesis to change C61 back to a serine and to introduce a new cysteine at residue 128 (S128C). Cells were harvested by centrifugation, resuspended in buffer A (25 mM Tris-HCl, pH 8.0 (25 °C), 500 mM NaCl, 50 mM ammonium acetate, 10 % glycerol) and lysed by sonication. After centrifugation, the supernatant was loaded onto a 2 ml Ni-NTA column (*Qiagen*), equilibrated with buffer A. After washing, bound protein was eluted with 250 mM imidazole in buffer A. The protein was then purified by a heparin sepharose column (HiTrap Heparin HP, 5 ml, *GE Healthcare*) using buffer A and a linear gradient from 250 mM to 1 M NaCl. After concentration, TBP core was further purified by application to a Superose 6 HR gel filtration column (*GE Healthcare*) equilibrated with Pol II buffer (50 mM HEPES pH 7.3, 40 mM ammonium sulphate, 10 μ M ZnCl₂, 20 % glycerol, 2 mM dithiothreitol (DTT)). Mutant TBP-S128C was directly subjected to dye labeling using a 10-15 fold molar excess of Alexa647-C2-Maleimide (*Molecular Probes*) in buffer A containing 600 mM NaCl at room temperature for 1 h. Free dye was removed using G-50 spin-columns (*GE Healthcare*) with assembly buffer (50 mM HEPES, 40 mM ammonium sulfate, 10 μ M ZnCl₂, 20 % glycerol, 10 mM DTT).

TFIIB

S. cerevisiae TFIIB with a C-terminal His-tag was expressed for 15 h at 20 °C in *E. coli* BL21 (DE3) RIL. Cells were harvested by centrifugation, resuspended in buffer B (50 mM Tris-HCl, pH 7.5 (25 °C), 500 mM NaCl, 10 mM imidazole, 5 % glycerol, 20 μ M PMSF, 5 mM β -mercaptoethanol) and lysed by sonication. After centrifugation, the supernatant was loaded onto a 2 ml Ni-NTA column (*Qiagen*), equilibrated with buffer B. After washing, bound protein was eluted with buffer B containing 1 M imidazole. The protein was diluted 1:3 with buffer C (50 mM Tris-HCl pH 7.5 (25 °C), 5 mM DTT) and was further purified by cation exchange chromatography (MonoS 5/50 GL, *GE Healthcare*) using buffer C and a linear gradient from 50 mM to 1 M NaCl. After concentration, a final gel filtration purification step was performed using a Superose 6 HR column (*GE Healthcare*) equilibrated with Pol II buffer containing 5 % glycerol.

TFIIB-ybbR

TFIIB-ybbR was expressed and purified as described above for wt TFIIB. Expression and purification of Sfp (plasmid gift from C. Kaiser) and synthesis of an Alexa647-CoA conjugate from Coenzyme A (*Sigma Aldrich*) and Alexa647-C2-Maleimide (*Molecular Probes*) was done as described [187]. TFIIB labeling was performed at room temperature for 30 min using 5 μ M TFIIB-ybbR, 1 μ M Sfp and 10-20 μ M Alexa647-CoA. Free dye was removed using Bio-Spin 6 columns (*Biorad*) equilibrated with assembly buffer containing 5 % glycerol (50 mM HEPES, 40 mM ammonium sulfate, 10 μ M ZnCl₂, 5 % glycerol, 10 mM DTT).

TFIIF

TFIIF subunit tfg1 from *S. mikatae* as well as TFIIF subunit tfg2 from *S. cerevisiae* were co-expressed recombinantly in *E. coli* and purified as described previously [248].

3.6.2. Nucleic Acid scaffolds

Nucleic acids scaffolds used to build OCs for smFRET measurements were constructed from 55 nt long DNA oligomers containing an 11 nt mismatch region around the TSS from register (+2) to (-9) (Figure 3.3.1A) and a TATA-box starting at register (-30). For surface immobilization of the OCs in the measurement chamber, the nontemplate DNA strand had biotin attached at the 5'-end via a C6-amino linker. DNA oligomers were annealed as described [15].

The FRET network used for NPS analysis of the OC is schematically represented in Figure 3.3.8: Template DNA strands were labeled at register (-10), (+3), (+7) or (+12) with Tamra (donor) or Alexa647 (acceptor) and used as satellites in order to map the position of antenna dyes on the upstream nontemplate DNA at register -30, -23 or -18, on TBP or on TFIIB. For localization of the antennas on the upstream nontemplate DNA, Tamra labeled nontemplate DNA strands were combined with Alexa647 labeled template strands. For localization of the antennas on TBP and TFIIB (Alexa647 labeled), Tamra labeled template strands were annealed with unlabeled nontemplate DNA strands. For control measurements, bubbles containing 17 nt RNA with the sequence 5'-AUGCAUAAAGACCAGGC-3' were used and annealed as described [15]. All DNA and RNA strands were purchased from *IBA* (Göttingen, Germany).

3.6.3. Open and elongation complex assembly

S. cerevisiae Pol II-TFIIF as well as 10- and 12-subunit *S. cerevisiae* Pol II were expressed and purified as described [70]. For OC assembly, TFIIB (60 pmol) was first added to the nucleic acid scaffold (15 pmol), followed by TBP (60 pmol) and the Pol II-TFIIF complex (10 pmol), yielding OCs in a final volume of 50 μ l assembly buffer (50 mM HEPES, 40 mM ammonium sulfate, 10 μ M ZnCl₂, 5 % glycerol) containing 10 mM DTT. The mixture

3 Dynamic architecture of a minimal Pol II open promoter complex

was incubated with slight agitation at room temperature for 1 h, followed by gelfiltration purification using a Superose 6 PC3.2 column (*GE Healthcare*). The fractions containing the complete OC were combined and diluted 100-1000 times in assembly buffer for sm-FRET measurements. In some experiments, labeled Rpb7-C150 was used as a satellite. Expression and purification of the single-cysteine mutant of Rpb4/7 as well as labeling of Rpb7-C150 was performed as described [15]. Labeled Rpb4/7 was then added in 15-20 fold molar excess to Pol II-TFIIF and exchange of unlabeled for labeled Rpb4/7 was allowed for 10-14 h at 4 °C. Subsequently, OC assembly was conducted as described above. OCs without TFIIF were assembled using the same assembly procedure. For OCs without TFIIF but with labeled Rpb4/7, 10 subunit core Pol II was used and pre-assembled with a 5 fold molar excess of labeled Rpb4/7 for 15 min prior to OC assembly. ECs were prepared consisting of 12 subunit Pol II and a nucleic acid scaffold containing 17 nt RNA as described previously [21].

3.6.4. Electrophoretic mobility shift assay

Typically, in a total volume of 15 μ l assembly buffer containing 10 mM DTT, 20 pmol of TBP, 20 pmol of TFIIB-ybbR labeled with Alexa647 and 5 pmol of the Pol II-TFIIF complex were sequentially added to 5 pmol of Tamra-labeled nucleic acids and incubated for 1 h at 25 °C. The samples were loaded on a 6-20 % gradient native polyacrylamide gel (*Biorad*) in 1x TBE buffer. The gel was prerun at 80 V for 45 min and the pockets were rinsed thoroughly before sample loading. Electrophoresis was carried out at 4 °C at a voltage of 120 V for 2 h and fluorescent bands were visualized using a Typhoon Imager (*GE Healthcare*) at 532 nm and 647 nm.

3.6.5. *In vitro* transcription assay

In an *in vitro* transcription assay, *HIS4* promoter activity was tested using wild type (wt) nuclear extract (NE) from *S. cerevisiae* BY4741, as well as NE from a temperature-sensitive *S. cerevisiae* strain SHY245 carrying a point mutation in the TFIIB gene SUA7. To this temperature-sensitive TFIIB NE, either protein storage buffer (negative control) or 5 pmol of wt TFIIB, TFIIB-(122)ybbR or TFIIB-(C-term)ybbR was added prior to start of *in vitro* transcription. NEs were prepared from 3 l of either BY4741 or SHY245 yeast culture as described previously [249]. *In vitro* transcription and analysis by primer extension was performed as described [249].

3.6.6. smFRET measurements: Experimental setup and data analysis

The homebuilt TIRF microscope setup used for all smFRET experiments as well as the data acquisition, processing and analysis procedures are discussed in sections 2.2.1 and 2.3. In all experiments, *sif*-movies with a total duration of 40 s and an integration time of 100 ms per frame were recorded and the resulting data was analyzed with the

custom-written MATLAB (*The MathWorks*) software *SM FRET V5.7* [196]. The resulting histograms were computed from every frame of the time traces smoothed by a 5 point moving average (framewise histogram). The histograms were fitted with one or more Gaussian distributions and the mean FRET efficiency as well as its standard error was determined from the fit. The overall error of the FRET efficiency was estimated by performing error propagation using the standard error from the fit and systematic errors and yielded to about 2 % and in a few cases to 3 % when only a low number of molecules with a broad distribution of FRET efficiencies could be obtained. The mean FRET efficiency and its overall error were used for further analysis by NPS [20, 22]. An overview of all smFRET data is given in Table A.1.1 and framewise FRET efficiency histograms are shown in Figures 3.3.4, A.1.1, A.1.2 and A.1.3.

3.6.7. Hidden Markov Model (HMM) analysis of dynamic smFRET time traces

Dynamic smFRET time traces were selected and HMM analysis was performed using a custom-written MATLAB (*The MathWorks*) software (section 2.4). The software is based on the MATLAB HMM-Toolbox by Kevin Murphy [214] and the used algorithm is similar to one previously described for HMM analysis of time-binned smFRET trajectories [212]. A two-state Hidden Markov process was assumed and an initial value for the width of states of 0.2 (20 % FRET) was chosen, as the Gaussian fits of the experimental FRET efficiency histograms are on average about 0.2 in width. HMM analysis was performed individually for all FRET traces (local analysis) instead of globally for the complete data set, because the examined OCs were expected to be to some extent heterogeneous in their composition. The presence of transcription factors TFIIB, TBP and TFIIF could not be inspected during the experiment and therefore incomplete OCs could not be discarded from the analysis. Since subpopulations of OCs lacking TBP, TFIIB or TFIIF are expected to show slightly different dynamics and FRET states than complete OCs, data acquired for the different complexes should not be analyzed globally. However, global analysis was also tested and yielded similar results. HMM analysis identified transitions between the two states in the time-binned smFRET time trajectories and all transitions were compiled in a transition density plot (TDP) (initial versus final FRET efficiency of every transition). The different transition populations in the TDP were selected and for each population, the cumulative events with a dwell time longer than a given time t were plotted against the time. A single exponential decay function was fitted to the cumulative distribution and the rate k of the transition could be extracted from the exponent.

3.6.8. Global Nano-positioning system analysis

Dye position priors

Satellite dye molecules were attached via flexible linkers to positions on the template DNA and on Pol II within OCs. The attachment points were assumed to be unchanged compared to their position within the Pol II elongation complex (EC), of which a structure has been solved by x-ray crystallography (pdb ID: 1Y1W) [51]. This assumption was validated by distance measurements discussed in section 3.3.2. Hence, the attachment points are known from the x-ray structure, whereas the precise location and the orientation of the dye molecules are not. For NPS analysis, we therefore calculated the volume that is sterically accessible to the dye molecules, given the point of attachment, the size of the dye molecule, the linker length and the diameter of the linker [20]. Briefly, the satellites were approximated by a sphere of a diameter $d_{dye} = 7 \text{ \AA}$, linked to template DNA and protein by flexible linkers with a diameter of $d_{linker} = 4.5 \text{ \AA}$ and a length of $L_{linker} = 13 \text{ \AA}$ (internal DNA labels) and 7 \AA (protein labels), respectively (Table 3.6.1). Additionally, we assumed vanishing probability density within the volume already occupied by the reference structure of the Pol II EC shrunk by 2 \AA to account for uncertainties in the x-ray structure. Each satellite position within the accessible volume was given equal probability (flat prior) [20]. The position and size of the computed satellite priors are displayed in Figure 3.3.1C. Antenna dye molecules were attached via flexible linkers to positions on the upstream nontemplate DNA (NT(-18)), the TATA box (NT(-23), NT(-30)), TBP (C128) and TFIIB ((C-term)ybbR and (122)ybbR). All attachment points were unknown in the coordinate system of the Pol II EC and were to be determined within by application of global NPS analysis. However, we used information available about the local macromolecule structure, given that this structure is not subject to conformational change when incorporated into an OC, and simulated accessible volumes relative to these structures in a similar way like for the satellite dyes (Figure 3.3.1B). By displacing the structures, also the antenna dyes were moved to other positions so that the structures served as frames of reference to each antenna dye. To simulate the accessible volume of the antenna on the upstream nontemplate DNA, we used the structure of a 12 nt B-DNA (pdb ID: 1BNA) [231]. For the TATA box and TBP positions, the simulation was based on a structure of a B-core-TBP-TATA-element complex (pdb ID: 1VOL without TFIIB) [55]. The final position and orientation of the antennas was calculated from their position and orientation relative to the local structure and from the position and orientation of the reference frames attached roughly to the center of mass of each local structure. The volume accessible to the origin of the docked reference frames was set to a cube of $200 \text{ \AA} \times 290 \text{ \AA} \times 220 \text{ \AA}$ with the origin at position $(x,y,z) = (0, 105, -90) \text{ \AA}$, while their orientation remained completely free. The antennas attached to TFIIB were treated in a different way, because the conformation TFIIB adopts in the OC is likely to deviate from its conformation in the TFIIB-Pol II structure [42] or in the B-core-TBP-TATA structure [55]. Therefore, we restricted the TFIIB antenna priors to a volume not occupied by Pol II in the EC structure. To this end, we assumed zero probability density

within the volume occupied by Pol II shrunk by 2 Å to account for uncertainties in the x-ray structure, and equal probability densities elsewhere in a cubic box around Pol II with a side length of 300 Å. The downstream DNA in the OC was observed to adopt two conformations, the conformation expected from the EC structure and an alternate conformation (see section 3.3.3 for details). We examined the alternate conformation by global NPS. To this aim, we treated the dye molecules attached to the downstream DNA as antennas with unknown position in combination with the FRET efficiencies of the side peaks present in all histograms of measurements between sites on the downstream DNA and sites on the upstream DNA or Rpb7-C150. Yet, for the main peaks in the FRET efficiency histograms, the dye molecules on the downstream DNA served as satellites with known positions. No reference frame was assigned to the antenna dyes on the downstream DNA because the DNA conformation is likely to deviate from the structure of a straight B-DNA in the alternate conformation, especially close to the melted region at register (+3). Therefore, we restricted the downstream DNA antenna priors to a volume not occupied by Pol II in the EC structure [51], in a similar way as for the TFIIB antenna priors.

Table 3.6.1.: Prior information of satellites and antennas. For each satellite the coordinates x,y and z of the dye attachment point are given in Angstroem in the Pol II coordinate system. Further, for each antenna assigned to a reference frame, the atom number, atom name and residue number of the attachment point is given in the reference frame. Additionally, the parameters of the flexible linker and the dye used for modeling the prior are listed: L_{linker} : length of flexible linker, d_{linker} : diameter of the linker, d_{dye} : diameter of the dye.

Satellites							
Dye position	Ref frame (pdb ID)	attachment point			L_{linker} [Å]	d_{linker} [Å]	d_{dye} [Å]
		x	y	z			
Rpb7-C150	1Y1W	117.283	30.008	-64.547	7	4.5	7
T-DNA(-10)	1Y1W	77.878	35.805	-16.395	13	4.5	7
T-DNA(+3)	1Y1W	95.713	59.792	-15.336	13	4.5	7
T-DNA(+7)	1Y1W	88.553	73.813	-19.986	13	4.5	7
T-DNA(+12)	1Y1W	94.929	87.763	-31.497	13	4.5	7

Antennas							
Dye position	Ref frame (pdb ID)	attachment point			L_{linker} [Å]	d_{linker} [Å]	d_{dye} [Å]
		chain	residue #	atom			
NT-DNA(-18)	1BNA moved to (0,0,0)	A	7	C6	13	4.5	7
NT-DNA(-23)	1VOL moved to (0,0,0)	C	11	C6	13	4.5	7
NT-DNA(-30)	1VOL moved to (0,0,0)	C	4	C6	13	4.5	7
TBP-C128	1VOL moved to (0,0,0)	A	86	CA	7	4.5	7

Determination of isotropic Förster distances and anisotropies

For each donor-acceptor pair the isotropic Förster radius R_0^{iso} was determined using standard procedures [250] (a detailed description can be found in the PhD thesis of J. Andrecka [200]). First, the quantum yield of the donor sample was determined by comparing its fluorescence to that of the standard Rhodamine 101 in ethanol ($Q = 100\%$). Second,

the overlap integral was calculated from the recorded donor fluorescence emission spectrum (emission wavelength 528-700 nm, excitation wavelength 523 nm) and the acceptor absorption spectrum (400-700 nm). Finally, a refractive index of $n = 1.35$ and $\kappa^2 = 2/3$ were used to calculate isotropic Förster distance R_0^{iso} . In order to account for uncertainties in the Förster distance due to orientation effects, we measured the steady-state fluorescence anisotropies $r_{i,\infty}$ of the donor and acceptor dyes for all attachment sites using a steady-state fluorescence spectrometer (*Edinburgh Instruments F900*) (Table 3.6.1). Both, isotropic Förster distances as well as fluorescence anisotropies were used in the global NPS analysis [22].

Global determination of the positions and orientations of antenna dyes and reference frames by NPS analysis

For details on global NPS analysis, see section 2.5.3. Briefly, Bayesian parameter estimation was applied using all measured average FRET efficiencies (with an error of 3 % for all measurements including TFIIB antennas and an error of 2 % for all other measurements), as well as all measured steady-state fluorescence anisotropies and isotropic Förster distances R_0^{iso} , to simultaneously infer position and orientation of all fluorophores within the "laboratory" coordinate system of Pol II. Moreover, the information of all fluorophores linked to a reference frame could be used to also position and orient the reference frame itself with respect to the Pol II EC coordinate system. As a result, we obtain as posterior the three-dimensional probability density function (PDF) of the positions and orientations of antenna and satellite dyes. As a feature of the docking NPS analysis, the marginal position PDF of any desired point (e.g. antenna attachment point) in the docked reference frames relative to the Pol II EC coordinate system can be extracted in addition to the marginal antenna dye position PDFs. The complete calculation was performed using a custom implementation of the nested sampling algorithm [224] based on Markov chain Monte-Carlo in C and MATLAB (*The MathWorks*). The NPS software is available free of charge at www.uni-ulm.de/nawi/nawi-biophys.html [22]. After NPS analysis, the marginal position PDFs of the antenna dyes and of the attachment points of the dyes within the docked reference frames relative to the Pol II EC coordinate system were exported as `XPLOR`- or `mrc`-files with a resolution of 6-8 Å. The credible volumes, i.e. the three-dimensional error bars, of the positions were finally displayed as isosurfaces in Chimera [151].

3.6.9. Modeling

The OC model was constructed using the molecular design software Moloc [251] and its stereochemistry was optimized. Small structural changes within the Pol II structure were introduced during energy-minimization of Rpb1 residues Asp268, Lys271, Lys317 and Rpb2 residues Arg430, Gln469, Lys470.

4. Single molecule FRET studies of nucleosome remodeling by Chd1

4.1. Abstract

In this chapter, I describe two smFRET based studies that aimed for shedding light on the mechanism of nucleosome remodeling by the remodeling enzyme Chd1.

In the first study, I explored the structural dynamics of nucleosomes during binding and remodeling by individual Chd1 remodelers in real-time by measuring smFRET between dye pairs attached to different positions of the nucleosomal DNA. The time resolved smFRET data revealed gradual and bidirectional translocation of nucleosomal DNA by continuously bound Chd1. The data allowed me to build a model of the remodeling mechanism applied by Chd1, which involves formation and propagation of a DNA loop.

In the second study, I used smFRET experiments, global NPS analysis and NPS docking to determine the molecular architecture of Chd1 bound to an end-positioned mono-nucleosome. I measured smFRET efficiencies between antenna dye molecules attached to the tandem chromodomains, the Snf2 ATPase domain and the DNA-binding domain of Chd1 and several satellite dye molecules attached to positions on the nucleosomal DNA within the nucleosome core particle. The data allowed me to build a preliminary model of the Chd1-nucleosome complex.

The work presented in this chapter was performed in collaboration with the laboratory of Professor Tom Owen-Hughes (University of Dundee, Scotland). The experiments for NPS positioning of Chd1 relative to the nucleosome were performed together with Monika Holzner, who joined this project as a new PhD student in April 2011.

4.2. Introduction

The packaging of DNA into chromatin represses essential DNA-associated processes, such as transcription, replication, DNA repair and recombination. Cells have evolved specialized ATP-dependent chromatin remodeling complexes equipped with the capability to slide (reposition) nucleosomes along DNA, thereby providing regulated access to the underlying DNA. Numerous biochemical studies have investigated nucleosome remodeling in bulk, however kinetic intermediates and pathways through which the nucleosome structure evolves during remodeling could not be resolved due to ensemble averaging. Moreover, standard high-resolution structural methods such as x-ray crystallography have been hindered by the large size, the oftentimes heterogeneous composition and the high flexibility of remodeler-nucleosome complexes. As a consequence, the molecular mechanism that allows remodelers to couple hydrolysis of ATP to the breakage of histone-DNA contacts without a full disruption of the nucleosome is poorly understood.

Single molecule experiments are ideally suited to probe the structural architecture and dynamics of remodeler-nucleosome complexes, since direct visualization of individual complexes in real-time is possible. Here, I used single molecule FRET and Nano-positioning system analysis to shed light on the mechanism of nucleosome remodeling by the chromatin remodeler Chd1. I chose to study Chd1 because of its simplicity and good molecular control being composed of a single polypeptide, because of the availability of crystal structures of individual domains and since, in contrast to ISWI-type remodelers, no data is available regarding its remodeling mechanism or its structure in complex with a nucleosome.

In the first study, I explored the structural dynamics of nucleosomes during binding and remodeling by individual Chd1 remodelers in real-time by using smFRET. Therefore, I used nucleosomal constructs labeled with FRET pairs at different positions of the nucleosomal DNA that allowed me to monitor conformational changes within the nucleosomal DNA during remodeling by Chd1. To obtain quantitative kinetic information of the remodeling dynamics, I analyzed dynamic FRET time trajectories using Hidden Markov Modeling analysis. The results allowed me to propose a model of the remodeling mechanism applied by Chd1, which involves formation and propagation of a DNA loop.

In the second study, I used smFRET experiments, global NPS analysis and NPS docking to determine the architecture of Chd1 bound to a nucleosome in the presence of the non-hydrolyzable ATP analog AMP-PNP. Therefore, I measured smFRET efficiencies between antenna dye molecules attached to the tandem chromodomains, the Snf2 ATPase domain and the DNA-binding domain of Chd1 and several satellite dye molecules attached to positions on the nucleosomal DNA within the nucleosome core particle. The resulting NPS position densities together with previous x-ray crystallographic information allowed me to build a preliminary model of the Chd1-nucleosome complex, which provides new insights into the mechanism of Chd1-catalyzed nucleosome remodeling.

4.3. Probing Chd1 induced nucleosome dynamics by smFRET

4.3.1. Results

Design of labeled nucleosomal templates for smFRET studies of Chd1 remodeling

I designed labeled nucleosomal templates amenable to smFRET investigation of the conformational changes a nucleosome undergoes during binding and remodeling by Chd1. The design was based on end-positioned mono-nucleosomes assembled on 200 bp long DNA containing the strong nucleosome positioning sequence 601 [111] and 47 and 6 bp long DNA linkers on the nucleosome entry and exit site, respectively (6-N-47). Chd1 was previously shown to reposition nucleosomes from the end towards the center of a short DNA template [128]. More precisely, 0-N-54 nucleosomes, similar to the 6-N-47 nucleosomes used in this study, were shown to be preferentially remodeled to 26-N-28 or 23-N-31 positions. Accordingly, a FRET dye pair had to be positioned onto the nucleosome such that repositioning by roughly 20 bp towards the longer linker could be observed as a reasonable change in FRET efficiency. Four different label positions were designed, of which two were on the nucleosomal DNA within the nucleosome core (F(+14)-Tamra and F(-71)-Alexa647) and two were on the extranucleosomal DNA either on the long linker (entry site, 12 bp outside the nucleosome core, R(-85)-Tamra/ -Alexa647) or on the short linker (exit site, 6 bp outside the nucleosome core, F(-term)-Alexa647) (Figure 4.3.1 and 4.6.1). Both labels within the nucleosome core are located on the same side of the nucleosome close to the exit site, but on different gyres.

From the four available labeling sites, three different nucleosomal templates were prepared containing the FRET dye pair combinations F(+14)-Tamra/ R(-85)-Alexa647 (F(+14)/ R(-85)), F(+14)-Tamra/ F(-71)-Alexa647 (F(+14)/ F(-71)) and F(-term)-Alexa647/ R(-85)-Tamra (F(-term)/ R(-85)). Nucleosomes with the dye combination F(+14)/R(-85) were expected to give initially a rather low FRET signal that changes into a high FRET signal upon remodeling by Chd1 (Figure 4.3.1). Moreover, this construct is sensitive to conformational changes of the long linker DNA relative to the nucleosome core and was expected to show a stepwise increase in FRET efficiency if remodeling occurred via the proposed loop-recapture model that involves sequential formation and propagation of a DNA bulge over the histone octamer surface (outlined in Figure 4.3.1, details in section 1.2.6, [252, 253, 170, 254, 255, 165]). In contrast, F(+14)/F(-71) nucleosomes are expected to exhibit high FRET prior to remodeling by Chd1 and low FRET thereafter. Besides, F(+14)/F(-71) nucleosomes are able to provide insights into dynamics of the short linker DNA relative to the nucleosome core. Finally F(-term)/R(-85) nucleosomes were designed to study the conformation of both DNA linkers relative to each other as well as a possible conformational change upon binding of Chd1.

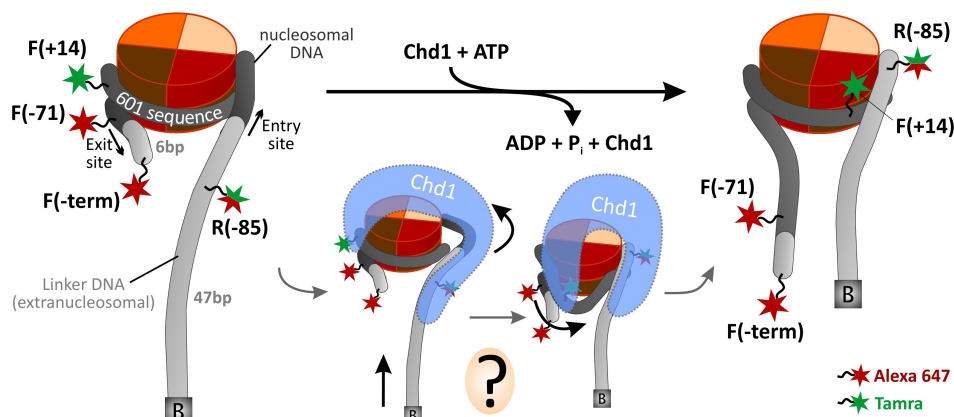


Figure 4.3.1.: Schematic illustration of labeled nucleosomes used for smFRET studies of Chd1 remodeling before (left) and after (right) the remodeling process. The hypothesized loop-capture mechanism is indicated below the arrow with Chd1 shown as semi-transparent blue object. Four labeling sites were designed (green stars, donor dye Tamra; red stars, acceptor dye Alexa647), of which two are situated on the nucleosomal DNA (F(+14), F(-71)) and the other two on the extranucleosomal DNA on either side (F(-term), exit site and R(-85), entry site). Nucleosomes containing the following FRET dye pair combinations were prepared: F(+14)-Tamra/ R(-85)-Alexa647, F(+14)-Tamra/F (-71)-Alexa647 or F(-term)-Alexa647/ R(-85)-Tamra.

Chd1 repositions double-labeled nucleosomes from the end towards the center of the 200 bp DNA under smFRET conditions

After preparing nucleosomal DNA and assembling labeled nucleosomes (see Experimental procedures 4.6), I performed a nucleosome remodeling assay to test whether fluorescent labels attached to the nucleosomal DNA impair Chd1 activity. Native gel analysis of the remodeling reactions (Figure 4.3.2) demonstrated that Chd1 is capable of repositioning labeled, end-positioned nucleosomes towards the center of the DNA and that the remodeler does not discriminate between labeled and unlabeled nucleosomes.

Next, I tested whether Chd1 is active on immobilized nucleosomes under smFRET measurement conditions. To this end, a smFRET experiment was performed with R(-85)/F(+14) nucleosomes, where FRET was recorded in the absence of remodeler as well as upon addition of truncated wild-type Chd1 (see section 4.6.9) to the nucleosomes in movies of 120 s duration. For each movie, a framewise FRET efficiency histogram was built from all detected FRET time trajectories. A notable change in FRET efficiency distribution could be detected upon addition of Chd1 to the nucleosomes (Figure 4.3.3). While nucleosomes alone showed a unimodal FRET efficiency distribution with a maximum at 45 % (green histogram), right after addition of Chd1 the distribution broadened and shifted towards higher FRET efficiencies, as expected for R(-85)/F(+14) nucleosomes. After 14 minutes (movie number seven), a well defined population had evolved at a FRET efficiency of 80 % comprising 75 % of all nucleosomes, whereas the population at 45 % FRET had become

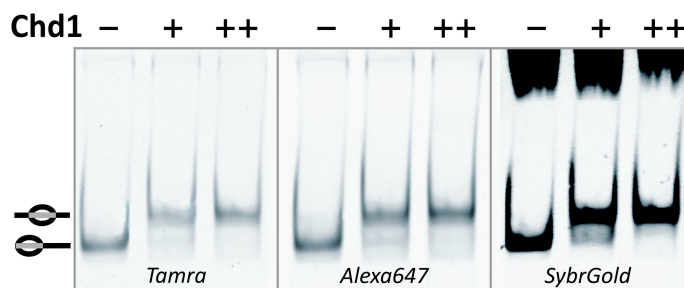


Figure 4.3.2.: Native gel analysis of remodeling of double-labeled mono-nucleosomes R(-85)-Alexa647/ F(+14)-Tamra catalyzed by Chd1. Green fluorescence of Tamra (left), red fluorescence of Alexa647 (middle) as well as the signal of the total DNA upon SYBR[®] Gold staining (lower panel) are shown. Chd1 is capable of sliding the end-positioned nucleosomes towards the center of the nucleosomal DNA, apparent as an upshift of the initial nucleosome band. The upper band in the SYBR[®] Gold stain image represents the unlabeled competitor DNA added to the reactions to stop the remodeling process by withdrawing Chd1 from the nucleosomal template.

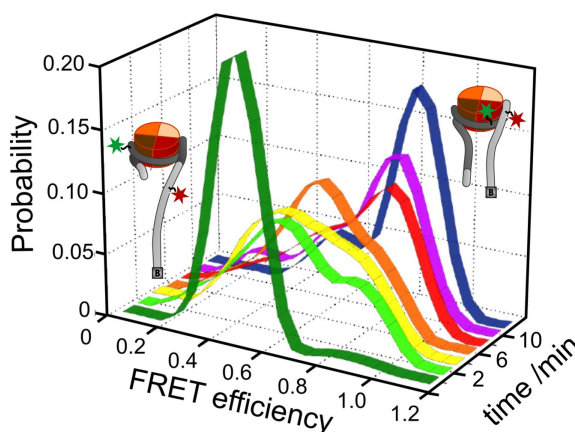


Figure 4.3.3.: Chd1 is able to reposition immobilized R(-85)/F(+14) nucleosomes under smFRET conditions. The temporal change in FRET efficiency distribution during remodeling by Chd1 is presented. For each histogram, the FRET traces of all selected nucleosomes of one movie (duration of movie, 120 s, number of nucleosomes per movie, 50-100) were compiled in a framewise manner. The first FRET efficiency histogram (green) shows the distribution prior to addition of Chd1, whereas all following histograms were recorded upon addition of 50 nM truncated wild-type Chd1 and 150 μ M ATP to the nucleosomes.

a subpopulation of only 25 % of nucleosomes. This distribution remained unchanged also for longer times. As a result, immobilized R(-85)/F(+14) nucleosomes are remodeled by Chd1 from the end towards the center of the 200 bp DNA.

In the following, I will go into detail of the smFRET observations during nucleosome remodeling by Chd1 in order to shed light on the underlying remodeling mechanism. First, I will focus on the nucleosome entry site with the long linker DNA by presenting smFRET experiments with R(-85)/F(+14) nucleosomes. Then I will examine Chd1 in-

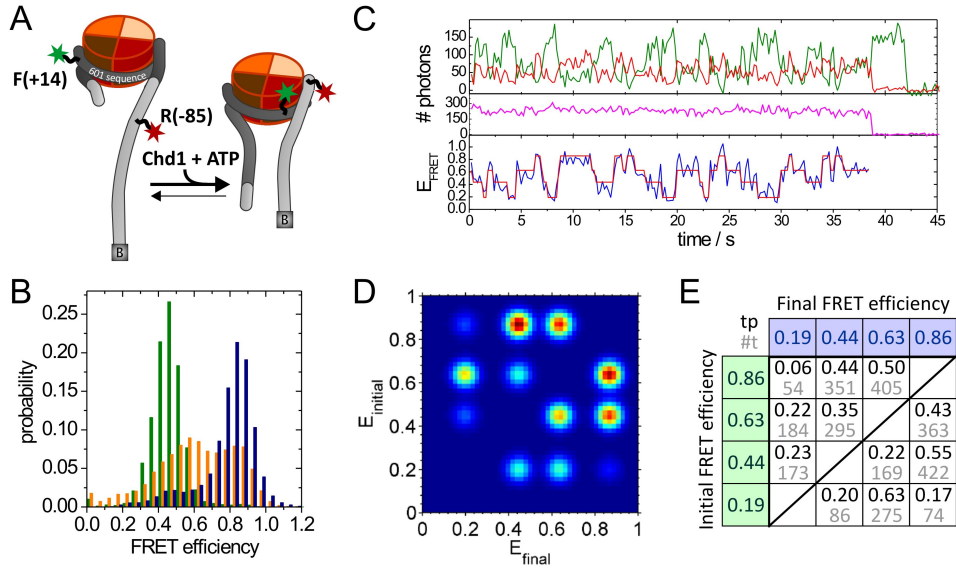


Figure 4.3.4.: Remodeling dynamics of R(-85)/F(+14) nucleosomes in the presence of 50 nM Chd1 and 150 μ M ATP investigated by smFRET. The combined results of three independent measurements (acquisition rate 10 Hz) are presented. (A) Cartoon illustrating the smFRET measurement of R(-85)/F(+14) nucleosomes before and after remodeling by Chd1. The change in distance between the two dyes upon remodeling is visible. (B) Framework FRET efficiency histogram of nucleosomes alone (green, 1326 molecules), nucleosomes showing a static FRET signal in the presence of Chd1 and ATP (blue, 2265 molecules) as well as nucleosomes showing dynamic transitions between different FRET states in the presence of Chd1 and ATP (orange, 394 molecules). All histograms are built from unsmoothed time trajectories. (C) Example of directly observed dynamic transitions between low-, intermediate- and high-FRET states of a single R(-85)/F(+14) nucleosome induced by Chd1 and ATP. Time trace of donor fluorescence (green, upper panel) and acceptor fluorescence upon green excitation (red, upper panel) are shown together with the acceptor fluorescence upon direct excitation at 637 nm (magenta, middle panel). In the lower panel, the computed FRET efficiency (blue) is presented together with the time dependent transition between different FRET states as identified by HMM analysis (red). (D) Transition density plot (TDP) for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from HMM analysis of a total of 378 dynamic traces showing 2850 transitions. (E) Matrix presenting transition probabilities for all detected transitions (tp , black numbers) together with the absolute number of detected transition ($\#t$, gray numbers). Note that the sum of all transition probabilities originating from one state (i.e. all probabilities in one line of the matrix) has to be 1. See Supplemental Figure A.2.4A for the transition rates extracted from a mono-exponential fit to the cumulative distribution of dwell times of each transition.

duced conformational changes with respect to the nucleosome exit site holding the short DNA linker by showing smFRET experiments with F(-71)/F(+14) nucleosomes. Finally, I will explore Chd1 induced dynamics of the two DNA linkers relative to each other by presenting measurements with F(-term)/R(-85) nucleosomes.

Chd1 induces processive, bidirectional dynamics in the nucleosomal DNA at the entry site during remodeling

To gain insights into remodeling dynamics, it is important to examine the real-time smFRET data, which can reveal information about the kinetics of transitions. Figure 4.3.4 summarizes the results for three independent smFRET experiments of remodeling of R(-85)/F(+14) nucleosomes in the presence of 50 nM Chd1 and 150 μ M ATP. The distance between the fluorophores at positions R(-85) and F(+14) is expected to show an overall decrease (possibly with intermediate increases of the distance) during repositioning of the nucleosome from the end towards the center of the DNA. Indeed, FRET substantially increased upon addition of Chd1 and ATP to the nucleosomes: while nucleosomes alone yielded a FRET efficiency distribution centered at 45 % FRET, in the presence of Chd1 and ATP the majority of nucleosomes emitting a static FRET signal (85 % of nucleosomes) showed a high FRET efficiency of 84 % (Figure 4.3.4B, green and blue histogram, respectively). In addition to this overall shift in FRET efficiency, a change in the kinetics of measured traces was observed upon addition of Chd1 and ATP to the nucleosomes. While nucleosomes alone emitted static FRET signals (99.4 % of all detected nucleosomes), in the presence of Chd1 and ATP a notable fraction of nucleosomes showed dynamics (Figure 4.3.4C, 70 % of nucleosomes in the first movie upon Chd1 addition, 17 % of nucleosomes overall). Compilation of all unsmoothed, dynamic time trajectories in a framewise FRET efficiency histogram (Figure 4.3.4B, orange) resulted in a broad multi-modal distribution, which was fitted with three Gaussians (fit not shown). An additional intermediate peak at 60 % FRET was visible beside the initial peak at 45 % FRET and the final peak at 84 % FRET of remodeled nucleosomes. The fact that the intermediate FRET population is visible in the histogram of dynamic nucleosomes but not in the histogram of static nucleosomes in the presence of Chd1 and ATP suggests that it originates from a short-lived remodeling intermediate only captured during the active, dynamic remodeling process.

More information about intermediate FRET states can be obtained by a direct analysis of dynamic smFRET traces. The traces revealed dynamic switching between multiple FRET states on the order of seconds. As expected for the proposed loop-recapture model, FRET changed in a step-wise manner, though not in a unidirectional way. To quantify the observed kinetics, I performed Hidden Markov Modeling (HMM) analysis (see section 2.4 and Experimental procedures 4.6.15), which provides information about the FRET values of individual states and about their dwell times and the kinetic rates of transitions between states. I tested different HMM models with different parameter sets and performed the analysis in a local or in a global way (see Supplemental Figures A.2.1 and A.2.2 for Model testing procedure). The data could be best described by global analysis assuming a four-state hidden Markov process with a fixed width of states ($\sigma = 0.02$), indicating one additional state compared to the three states detected by FRET efficiency histogram analysis.

4 Single molecule FRET studies of nucleosome remodeling by Chd1

Global HMM analysis of a total of 378 dynamic FRET traces undergoing 2850 transitions resulted in a transition density plot (TDP, Figure 4.3.4D) showing 12 distinct transition populations. Four different states were identified at the FRET efficiencies 19 %, 44 %, 63 % and 86 % and all possible transitions occurred in between these states. It came as a surprise that even though the remodeling process was found to have a well defined final state demonstrated by a single peak at 84 % in the FRET efficiency distribution of static nucleosomes, the observed dynamics in FRET did not show unidirectionality. From the TDP, the transition probabilities tp of all transitions could be extracted as well as the absolute occurrence of each transition as presented in the Table in Figure 4.3.4E. From the initial state at 86 % FRET, the final states at 63 % and 44 % FRET are mainly populated and from the initial state at 63 % FRET, transitions to 44 % FRET and to 86 % FRET are equally probable. The transition to the respective third state is less frequently observed in both cases. Moreover, from the initial state at 44 % FRET, the transition to 86 % FRET was most frequently observed with a probability of 55 %. This indicates a cycling between those three states. The fourth state at 19 % FRET is most frequently populated from the initial states at 44 % and 63 % FRET and in turn, the 19 % FRET state most frequently (probability of 63 %) transits to the state at 63 % FRET. Transitions to and from the lowest FRET state at 19 % occur approximately half as often as transitions between the three higher FRET states, which indicates that this low FRET state lies off the pathway between the other three states. Moreover, since the left upper half of the TDP above the diagonal does not present the mirror image of the right lower half below the diagonal, the pathway a nucleosome takes to evolve through the different FRET states during remodeling cannot be a linear sequence of the different states.

In addition to the transition probabilities, cumulative distributions for the dwell times of all transitions could be extracted from the TDP and fitted by mono-exponential decays (see Supplemental Figure A.2.3), which yielded the rates k of transitions (Supplemental Figure A.2.4A). All rates are on the time scale of hundreds of milliseconds to seconds ranging from $0.58 \pm 0.06 \text{ s}^{-1}$ for the 63 % \rightarrow 44 % transition to $1.29 \pm 0.08 \text{ s}^{-1}$ for the 19 % \rightarrow 63 % transition.¹ Kinetic rates from low FRET states (19 % and 44 % FRET) to the final FRET state at 86 % are higher than the reverse rates with the rate $k_{0.19 \rightarrow 0.86}$ being the highest of all rates (Figure 4.3.4E). This indicates that nucleosomes with the FRET efficiency of 86 % accumulate due to a faster movement towards this state. It has to be noted, that information about the dissociation rate of Chd1 from the nucleosome in the different remodeling states cannot be obtained by the used smFRET assay. The well defined end state of remodeling might be the result of Chd1 having the highest dissociation rate from nucleosomes in the conformation underlying the 86 % FRET state.

¹Note that the transition 19 % \rightarrow 86 % shows an even higher rate of 2.39 s^{-1} , however this rate has a large error due to low statistics and is therefore not reliable.

Remodeling dynamics of nucleosomal DNA are independent on Chd1 concentration

In the next experiment, I addressed the question whether the FRET dynamics observed during remodeling depend on Chd1 concentration. To this end, I conducted remodeling experiments with R(-85)/F(+14) nucleosomes and 150 μ M ATP, but this time with 5 nM instead of 50 nM Chd1 (Figure 4.3.5A). A similar remodeling result as in experiments with 50 nM Chd1 was obtained as can be seen in the FRET efficiency histograms of static and dynamic nucleosomes in the presence of Chd1 and ATP (Figure 4.3.5B compared to Figure 4.3.4B). The peak at a FRET efficiency of 45 % almost completely shifted to a higher FRET efficiency of 75 % (slightly lower FRET than in the experiment with 50 nM Chd1) upon addition of Chd1 and ATP (Figure 4.3.5B, green and blue histogram, respectively). Moreover, a notable fraction of dynamic nucleosomes (13 % of all detected nucleosomes) was observed yielding a broad, multimodal FRET efficiency distribution (Figure 4.3.5B, orange histogram). Further, the overall time required for the shift in FRET efficiency from \sim 45 % to \sim 80 % agreed well with the required time in the measurement with 50 nM remodeler (about 15 min). Global HMM analysis (same model parameters as for analysis of remodeling experiments with 50 nM Chd1: 4 states, fixed $\sigma = 0.02$) of a total of 242 dynamic traces showing 809 transitions yielded a TDP that shows four different FRET states at the FRET efficiencies 14 %, 43 %, 60 % and 84 % (Figure 4.3.5D), similar to the result of the 50 nM Chd1 experiments. In addition, the transition probabilities obtained from the TDP (Figure 4.3.5E) and the transition rates (Supplemental Figure A.2.4B) extracted from the mono-exponential decay fit to the cumulative distribution of dwell times of each transition (not shown) were consistent for both Chd1 concentrations. This shows that the observed FRET dynamics during Chd1 remodeling are independent of the concentration of Chd1, at least in the range between 5-50 nM Chd1, suggesting that the observed bidirectional remodeling dynamics were accomplished by continuously bound Chd1 without requiring dissociation and rebinding of Chd1 from solution.

Chd1 induces conformational changes in the long linker DNA in the absence of ATP

In order to separate dynamics that require the energy of ATP-hydrolysis from those that do not depend on ATP, I performed experiments with R(-85)/F(+14) nucleosomes in the presence of Chd1 alone as well as in the presence of Chd1 and the non-hydrolyzable ATP analog AMP-PNP. The results are shown in Figure 4.3.6.

In the presence of Chd1 alone or with AMP-PNP, R(-85)/F(+14) nucleosomes did not undergo an overall conformational change since the FRET efficiency histogram of static nucleosomes before and after addition of Chd1 was similar with a maximum at 42-45 % FRET (Figure 4.3.6B, green and blue histograms, respectively). Only for nucleosomes with Chd1 and AMP-PNP, an additional side peak comprising 18 % of the data was visible at a low FRET efficiency of 10 %. However, like in the experiment with Chd1 and ATP, a notable fraction of dynamic nucleosomes appeared in the presence of Chd1

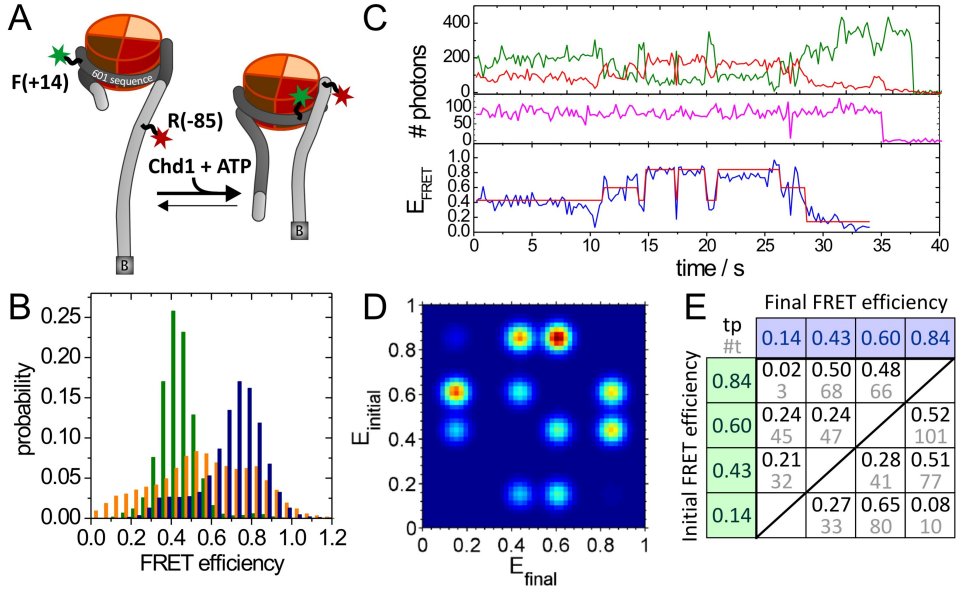


Figure 4.3.5.: Dynamics of R(-85)/F(+14) nucleosomes in the presence of 5 nM Chd1 and 150 μ M ATP investigated by smFRET. The combined results of three independent measurements are presented. (A) Cartoon illustrating the smFRET measurement of R(-85)/F(+14) nucleosomes before and after remodeling by Chd1. The change in distance between the two dyes upon remodeling is visible. (B) Framework FRET efficiency histogram of nucleosomes alone (green, 928 molecules), nucleosomes showing a static FRET signal in the presence of Chd1 and ATP (blue, 1681 molecules) as well as nucleosomes showing dynamic transitions between different FRET states in the presence of Chd1 and ATP (orange, 242 molecules). All histograms contain unsmoothed time trajectories. (C) Example of directly observed dynamic transitions between low-, intermediate- and high-FRET states of a single R(-85)/F(+14) nucleosome induced by Chd1 and ATP. Time trace of donor fluorescence (green, upper panel) and acceptor fluorescence upon green excitation at 637 nm (magenta, middle panel). In the lower panel, the computed FRET efficiency (blue) is presented together with the time dependent transitioning between different FRET states as identified by HMM analysis (red). (D) TDP for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from HMM analysis of a total of 242 dynamic traces showing 809 transitions. (E) Matrix presenting transition probabilities for all detected transitions (tp , black numbers) together with the absolute occurrence of each transition ($\#t$, gray numbers). Note that the sum of all transition probabilities originating from one state (i.e. all probabilities in one line of the matrix) has to be 1. See Supplemental Figure A.2.4B for the transition rates extracted from a mono-exponential fit to the cumulative distribution of dwell times of each transition.

4.3 Probing Chd1 induced nucleosome dynamics by smFRET

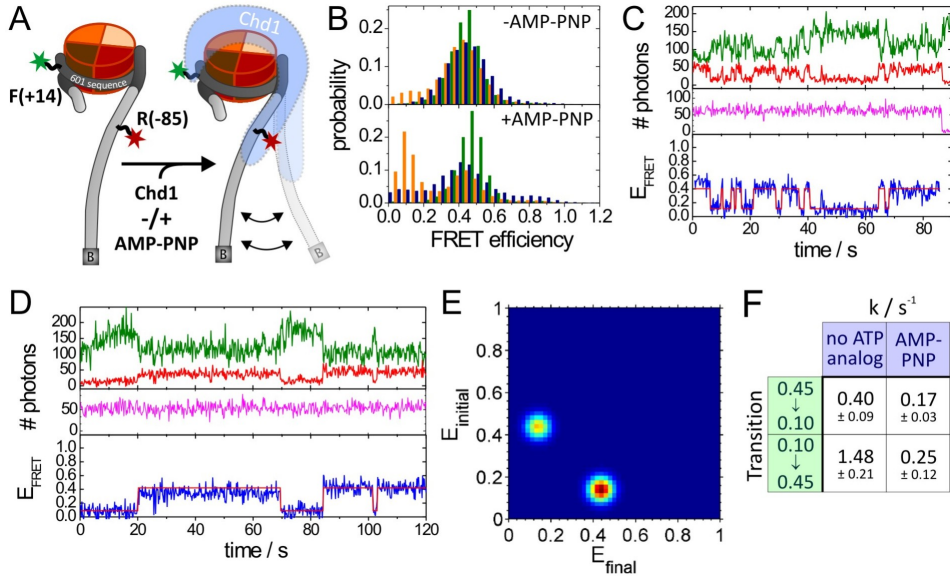


Figure 4.3.6.: Dynamics of R(-85)/F(+14) nucleosomes in the presence of 50 nM Chd1 alone or with 150 μ M AMP-PNP explored by smFRET. The combined results of three independent measurements for each condition are presented. (A) Cartoon illustrating the smFRET measurement of R(-85)/F(+14) nucleosomes before and after addition of Chd1 with and without AMP-PNP. Observed dynamics of the long linker DNA (see panel C-F) in the presence of Chd1 (blue shape) are indicated. (B) Framework FRET efficiency histograms of nucleosomes alone (green), nucleosomes showing static FRET in the presence of Chd1 (blue) as well as nucleosomes showing dynamic FRET traces in the presence of Chd1 (orange). The upper panel shows data of the measurements without ATP analog (green: 720 molecules, blue: 779 molecules, orange: 184 molecules) and the lower panel shows data of the measurements with AMP-PNP (green: 1415 molecules, blue: 624 molecules, orange: 188 molecules). Each histogram is built from unsmoothed time traces. (C) Exemplary dynamic FRET time trace of a single R(-85)/F(+14) nucleosome in the presence of Chd1 without any ATP analog. Time trace of donor fluorescence (green, upper panel) and acceptor fluorescence upon green excitation (red, upper panel) are shown together with the acceptor fluorescence upon direct excitation at 637 nm (magenta, middle panel). In the lower panel, the computed FRET efficiency (blue) is presented together with the time dependent transitioning between different FRET states as identified by HMM analysis (red). (D) Exemplary dynamic FRET time trace of a single R(-85)/F(+14) nucleosome in the presence of Chd1 and AMP-PNP. Same color coding as in panel C. (E) TDP for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from HMM analysis of a total of 184 dynamic traces of nucleosomes in the presence of Chd1 alone showing 791 transitions. A similar TDP was obtained for 188 dynamic traces of R(-85)/F(+14) nucleosomes with Chd1 and AMP-PNP showing 189 transitions. (F) Table presenting transition rates and their standard deviations in s^{-1} of the two detected transitions for the two different conditions: 50 nM Chd1 without ATP analog and with 150 μ M AMP-PNP. Each rate was extracted from a mono-exponential fit to the cumulative distribution of dwell times of each transition. The error is given as standard deviation of the rate from randomly chosen subsets of molecules undergoing the same transition.

without and with AMP-PNP. The framewise FRET efficiency histogram for the dynamic fraction of nucleosomes in the absence of any ATP analog (Figure 4.3.6B, upper panel, orange) showed a major peak centered at 45 % FRET consistent with the peak of static nucleosomes, but in addition showed a subpopulation at 10 % with a probability of 14 %. In the presence of AMP-PNP, this additional population had a probability of 55 % and hence a stronger peak than the one at 45 % FRET. Exemplary dynamic FRET time trajectories are given in Figure 4.3.6C and D in the presence of Chd1 without ATP analog and with AMP-PNP, respectively. Both trajectories reveal dynamic switching of the nucleosomes between the initial FRET state at 45 % and the state at 10 % FRET. This low FRET state indicates an alternative conformation of the nucleosome, in which the long linker DNA and hence the fluorophore at position R(-85) is located further away from the opposite linker side with the fluorophore at position (+14). The data suggests that Chd1 binds to the nucleosome and induces this conformational change in the long linker DNA (see caption in Figure 4.3.6A).

Comparison of the time trajectories without and with AMP-PNP suggests that switching between the two states occurs faster in the absence of AMP-PNP. For quantification of the kinetics, global HMM analysis was performed of a total of 184 traces showing 791 transitions for nucleosomes with Chd1 alone and of a total of 188 dynamic traces showing 189 transitions for nucleosomes with Chd1 and AMP-PNP. The resulting TDPs were similar for both conditions showing two distinct transitions from 45 % FRET to 10 % FRET and vice versa (Figure 4.3.6E), yet the transition rates as well as their ratios deviated.² In the presence of AMP-PNP, the 45 % \rightarrow 10 % transition occurred at a rate of $k_{0.45 \rightarrow 0.10} = 0.17 \pm 0.03 \text{ s}^{-1}$ and the reverse transition was 1.5 times faster ($k_{0.10 \rightarrow 0.45} = 0.25 \pm 0.12 \text{ s}^{-1}$). In contrast, in the absence of any ATP analog, the 45 % \rightarrow 10 % transition was 2.4 times faster ($k_{0.10 \rightarrow 0.45} = 0.40 \pm 0.09 \text{ s}^{-1}$) than the same transition in the presence of AMP-PNP and the transition back was even 5.9 times faster ($1.48 \pm 0.21 \text{ s}^{-1}$). Hence, the ratio of rates $k_{0.10 \rightarrow 0.45}$ to $k_{0.45 \rightarrow 0.10}$ increased from 1.5 in the presence of AMP-PNP to 3.7 in the absence of AMP-PNP consistent with the different relative probabilities of the two FRET states for both conditions in the FRET efficiency histograms of dynamic traces (Figure 4.3.6B). This result shows that the dynamic switching of nucleosomes between different FRET states is dependent on AMP-PNP, which slows down the dynamics. Moreover, the alternative conformation of nucleosomes with the long linker DNA further away from the dyad is stabilized in the presence of the non-hydrolyzable nucleotide AMP-PNP. This explains the presence of the side peak at 10 % FRET in the FRET efficiency histogram of static nucleosomes with Chd1 and AMP-PNP in contrast to the condition without ATP-analog.

²Note that the intensities of both transition populations in the TDP are not identical, as would be expected for a system in equilibrium. The fact that the transition from the higher FRET state back to the lower FRET state is less often populated indicates, that bleaching occurs with a higher probability from the higher FRET state. This also shows, that the examined system is not in equilibrium.

The nucleosome dynamics observed in the presence of Chd1 alone or with AMP-PNP are assumed to depict the first step of the complete remodeling process prior to ATP hydrolysis. Indeed, similar transitions between an initial state at 43-44 % FRET and a low FRET state at 14-19 % FRET were observed as part of the remodeling process of R(-85)/F(+14) nucleosomes (Figure 4.3.4D and E and Figure 4.3.5D and E). In the presence of 150 μ M ATP, the rates of the forward and backward transitions are similar, $k_{0.44 \rightarrow 0.19} = 0.93 \pm 0.14 \text{ s}^{-1}$ and $k_{0.19 \rightarrow 0.44} = 0.89 \pm 0.24 \text{ s}^{-1}$ (see Figure A.2.4), and higher than the rates measured for nucleosomes in the presence of Chd1 and AMP-PNP, presumably because ATP hydrolysis drives the nucleosomes out of the 19 % FRET state. Compared to the rates measured for nucleosomes with Chd1 but without ATP analog, the rate for the transition $0.44 \rightarrow 0.19$ in the presence of ATP is twice as high, however the rate for the reverse transition is two times slower in the presence of ATP. When ATP is available, a different transition out of the 19 % FRET state than the transition back to 44 % FRET, namely the transition to a final FRET efficiency of approximately 60 % FRET is preferred, since it has a faster rate than the transition back to 44 % FRET ($k_{0.19 \rightarrow 0.60} = 1.29 \text{ s}^{-1} / 1.34 \text{ s}^{-1}$ for 50 nM/ 5 nM Chd1).

Chd1 stably binds nucleosomes even in the absence of ATP

It remains unclear, whether the above described nucleosome dynamics observed in the presence of Chd1 with or without AMP-PNP correlate with Chd1 binding and dissociation events or whether Chd1 is constantly bound to the nucleosome during the switching between different FRET states. To test the stability of Chd1 binding to nucleosomes, the following experiment was performed (see schematic in Figure 4.3.7A): R(-85)/F(+14) nucleosomes were first incubated with 50 nM Chd1 for 15 min before unbound Chd1 was removed from solution by thorough rinsing (400 μ l remodeling buffer at a flow rate of 100 μ l/min) with remodeling buffer. Finally, ATP was added to the nucleosomes and smFRET measurements were recorded. If Chd1 was stably bound to the end-positioned nucleosomes, one would expect remodeling to occur upon addition of ATP. However, if Chd1 dissociated from nucleosomes within the time frame of 4 min (duration of rinsing), no change in FRET efficiency would be observed, because, once dissociated into the chamber volume, the probability of rebinding is extremely low.

Interestingly, active remodeling was observed upon addition of ATP to nucleosomes with prebound Chd1. While nucleosomes in the presence of Chd1 without nucleotide showed a single FRET peak at about 45 % FRET as expected from previous experiments, after removal of Chd1 from solution and addition of ATP, 53% of nucleosomes showed a FRET efficiency of 82 %, characteristic for remodeled R(-85)/F(+14) nucleosomes (Figure 4.3.7B, green and blue histogram, respectively). In addition, a total fraction of 17 % of nucleosomes showed dynamic FRET trajectories, which yielded a multimodal FRET efficiency distribution with peaks at 9 %, 39 %, 60 % and 84 % FRET (Figure 4.3.7B,

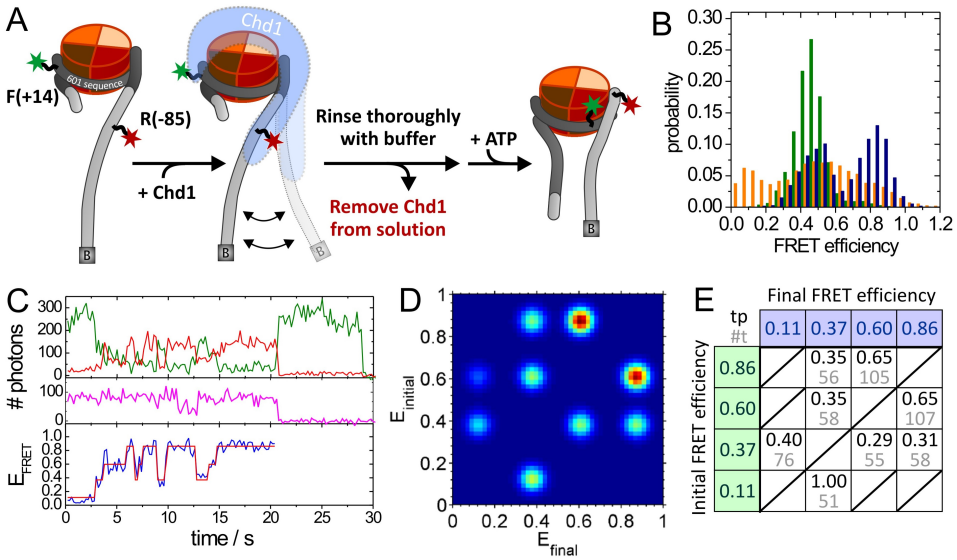


Figure 4.3.7.: Chd1 stably binds R(-85)/F(+14) nucleosomes even in the absence of nucleotide. (A) Schematic illustration of the experiment outline: immobilized R(-85)/F(+14) nucleosomes were first incubated with 50 nM Chd1 in the absence of nucleotide. Subsequently, the measurement chamber was thoroughly rinsed to remove unbound Chd1. Finally, remodeling buffer supplemented with 150 μ M ATP was added to the nucleosomes and smFRET data was acquired. (B) Framework FRET efficiency histograms are given for nucleosomes in the presence of Chd1 (green, 350 molecules) and of static (blue, 381 molecules) and dynamic (orange, 65 molecules) nucleosomes upon removal of Chd1 from solution and addition of ATP. All histograms are built from unsmoothed time trajectories. (C) Exemplary dynamic FRET time trace of a single R(-85)/F(+14) nucleosome with prebound Chd1 and in the presence of 150 μ M ATP. The same color coding is used as in the previous figures. (D) TDP for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from HMM analysis of a total of 106 dynamic traces showing 590 transitions (combined dynamic traces of two independent measurements). Eight transition populations are visible between four different states at 11 %, 37 %, 60 % and 86 % FRET. (E) Table presenting transition probabilities for all detected transitions. Note that the sum of all transition probabilities originating from one state (i.e. all probabilities in one line of the matrix) has to be 1. See Supplemental Figure A.2.4C for the transition rates extracted from a mono-exponential fit to the cumulative distribution of dwell times of each transition.

orange histogram, fit not shown, and C). The fact that remodeling was not complete can be attributed to loss of activity of Chd1 during exposure at room temperature, which was observed in all smFRET remodeling experiments as presented later on (Figure 4.3.11) as well as to a partial dissociation of Chd1 from the nucleosomes. Unlike in previous experiments, where fresh Chd1 was supplied together with ATP, in this experiment Chd1 had resided for roughly 15 min in the measurement chamber at room temperature before ATP was added for remodeling to begin and hence, once ATP was present, only a subset of Chd1 was still active and bound. HMM analysis was performed with a total of 106 dynamic traces (2 measurements combined) showing 590 transitions and the resulting TDP as well as the transition probability matrix are shown in Figure 4.3.7D and E. The

four detected FRET states at 11 %, 37 %, 60 % and 86 % FRET were consistent with the states found in previous experiments, yet only eight transition populations were detected compared to 12 populations in the experiment with 50 nM Chd1 constantly present in the measurement chamber. Transitions between the low FRET state at 11 % and the states at 60 % and 86 % FRET were missing. The transition probability matrix mostly agreed with previous experiments, only for the initial state at 37 % FRET, the transition to the 11 % FRET state was most probable, whereas in previous experiments, transitions to the FRET states at 60 % or 86 % FRET were more probable. Transition rates extracted from the mono-exponential decay fit to the cumulative distribution of dwell times (Figure A.2.4C) were generally higher than for measurements with Chd1 constantly present in solution, but also less precise due to relatively low statistics.

In summary, this experiment demonstrates that Chd1 has a strong commitment to nucleosomal templates and stably binds nucleosomes even in the absence of any ATP analog. Moreover, it shows that the observed processive back-and-forth movement observed during Chd1 remodeling is accomplished by stably bound Chd1 and does not require additional Chd1 from solution. Thus, dynamics of the long linker DNA observed in the presence of Chd1 alone or with AMP-PNP are the result of the action of constantly bound Chd1 that induces conformational changes in the nucleosome even in the absence of nucleotide. It will be shown later (section 4.4) in experiments with fluorescently labeled Chd1 and singly-labeled F(+14)-Tamra nucleosomes that the DNA-binding domain (DBD) of Chd1 shows similar dynamics relative to the nucleosome core as the long linker DNA. Hence, the Chd1 DBD is assumed to not only induce the dynamics in the long linker DNA, but bind the linker DNA and move together with it.

ATP dependency of Chd1 induced dynamics of nucleosomal DNA

In order to explore the ATP dependency of the observed dynamic switching between different FRET states, remodeling of R(-85)/F(+14) nucleosomes was performed with 50 nM Chd1 and 1 μ M instead of 150 μ M ATP. In addition, an ATP recovery system based on creatine phosphate and creatine phospho-kinase was used to avoid ATP depletion. A similar remodeling result and similar FRET dynamics were observed as in the experiments with 150 μ M ATP. Quantitative HMM analysis was performed with a total of 98 dynamic traces showing 396 transitions resulting in a TDP similar to the one obtained for remodeling with 150 μ M ATP (Figure 4.3.8A). Four different states were found at 16 %, 37 %, 54 % and 78 % FRET and transitions in between all states were detected consistent with the high ATP results. Transition probabilities were extracted from the data as presented in Figure 4.3.8B. Notably, the transitions between the states at 16 % and 37 % FRET were most highly populated, which is understandable since those transitions were shown earlier to be induced by Chd1 even in the absence of ATP. Beside this deviation, transition probabilities agreed well with results from high ATP experiments. In addition to transition probabilities, transition rates were extracted from the TDP (Figure 4.3.8C)

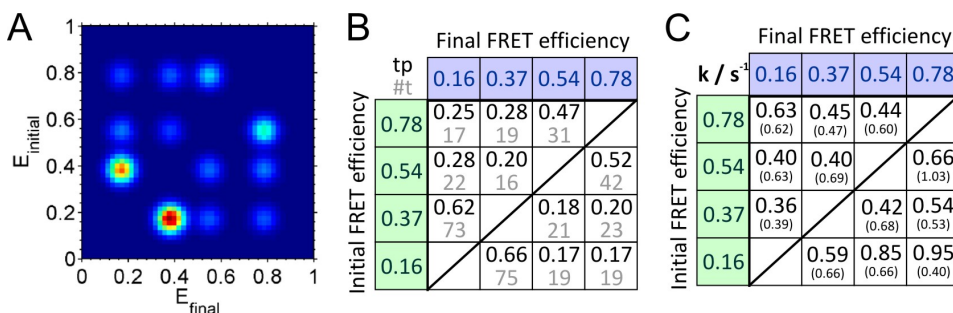


Figure 4.3.8.: Remodeling of R(-85)/F(+14) nucleosomes with 50 nM Chd1 and 1 μ M ATP. (A) TDP for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from HMM analysis of a total of 99 dynamic traces showing 396 transitions. 12 transition populations are visible between four different states at 16 %, 37 %, 54 % and 78 % FRET. (B) Matrix presenting transition probabilities for all detected transitions (tp , black numbers) together with the absolute occurrence of each transition ($\#t$, gray numbers). Note that the sum of all transition probabilities originating from one state (i.e. all probabilities in one line of the matrix) has to be 1. (C) Table presenting transition rates of all detected transitions. Each rate was extracted from a mono-exponential fit to the cumulative distribution of dwell times of each transition. In parenthesis, the ratios of the transition rates for remodeling with 1 μ M ATP and 150 μ M ATP are shown.

and were generally 1.5-2.5 times lower than the rates obtained in the experiment with 150 μ M ATP, suggesting that all transitions are dependent on the energy of ATP hydrolysis. Strangely, also the rates of transitions between 16 % and 37 % FRET were slower, which would not have been expected if those transitions are fully independent of ATP. However, these results have to be treated with caution since they are based on a single experiment and not enough switching events were observed for a reliable extraction of rates. The experiment should be repeated to solidify the drawn conclusions.

Processive, bidirectional remodeling dynamics also observed for nucleosomal DNA at the exit site

In all the above experiments with R(-85)/F(+14) nucleosomes, only one distance was monitored during the remodeling process between the long linker DNA and a position on the nucleosomal DNA within the nucleosome core 1.5 turns away from the dyad. An overall decrease in distance due to repositioning by Chd1 was observed with stepwise, dynamic distance changes during the active remodeling process. Since both labels are being repositioned during the remodeling event, a structural interpretation of the different FRET states is difficult. More distances spanning different parts of the nucleosome need to be investigated to allow for an interpretation of the intermediate nucleosome conformations underlying the different FRET states.

Therefore, I performed a set of smFRET experiments with Chd1 and F(-71)/F(+14) nucleosomes, which are sensitive to dynamics occurring on the nucleosome exit site holding the short DNA linker. The nucleosomes alone were expected to show a high FRET effi-

ciency that shifts to lower FRET efficiencies during repositioning towards the center of the DNA by Chd1 (bulk experiments, data not shown). In the first experiment, I explored remodeling in the presence of 50 nM Chd1 and 150 μ M ATP (Figure 4.3.9A). As expected, F(-71)/F(+14) nucleosomes alone showed a high FRET efficiency of 76-79 %, which is consistent with the distance expected from the crystal structure of a nucleosome assembled on the 601 sequence [256] (Figure 4.3.9B, green histogram). However the FRET efficiency histogram of static nucleosomes in the presence of Chd1 and ATP did not meet the expectations. Instead of the anticipated low FRET, the FRET efficiency distribution, which was fitted with four Gaussians (fit not shown), showed a major peak with a probability of 56 % centered at a high FRET efficiency of 90 %. Besides, subpopulations were visible at 11 % FRET (12 % probability), 53 % FRET (11 % probability) and the initial FRET efficiency of 75 % (20 % probability). The unexpected high FRET efficiency of remodeled nucleosomes might be the result of a conformation of the nucleosome, in which the linker DNA, which had been extended during the repositioning process, exits the nucleosome in a small angle and hence crosses the other linker on the nucleosome entry site (Figure 4.3.9A, right cartoon). Such conformation would bring the labels at position F(+14) and F(-71) close. Consistent with previous experiments, upon addition of Chd1 and ATP a total fraction of 38 % of nucleosomes were detected to show FRET dynamics. The FRET efficiency histogram of dynamic nucleosomes yielded a broad distribution spanning the whole FRET range, which was fitted with three Gaussians with maxima at 13 %, 38 % and 84 % FRET comprising 32 %, 40 % and 28 % of the data, respectively. Hence, even though the FRET efficiency of static remodeled nucleosomes is unexpectedly high, nucleosomes showing dynamic FRET traces during Chd1-mediated remodeling exhibit a substantial fraction of low FRET. The dynamic FRET time trajectories revealed dynamic, bidirectional switching between multiple FRET states on the order of hundreds of milliseconds to seconds (Figure 4.3.9C).

Time scale of remodeling dynamics

In order to test whether the acquisition rate of 10 Hz applied so far in all experiments was suitable to detect all remodeling dynamics, measurements with an acquisition rate of 30 Hz were made in addition to the 10 Hz measurements. Since ALEX was performed, the achieved time resolution in each channel was 200 ms in the case of 10 Hz and three times higher (66 ms) in the case of 30 Hz. The results for the measurements at a time resolution of 30 Hz are shown in Figure 4.3.9 with respect to the FRET efficiency histograms (panel D) and the dynamic FRET time trajectories (panel E). The FRET efficiency histograms of nucleosomes alone and of static and dynamic nucleosomes in the presence of Chd1 for the 30 Hz experiments were similar to those revealed from 10 Hz experiments, only that peaks were generally wider. Comparison of exemplary FRET time trajectories of the measurements at both time resolutions revealed similar dynamic, bidirectional switching between different FRET states, however the switching rate appeared substantially

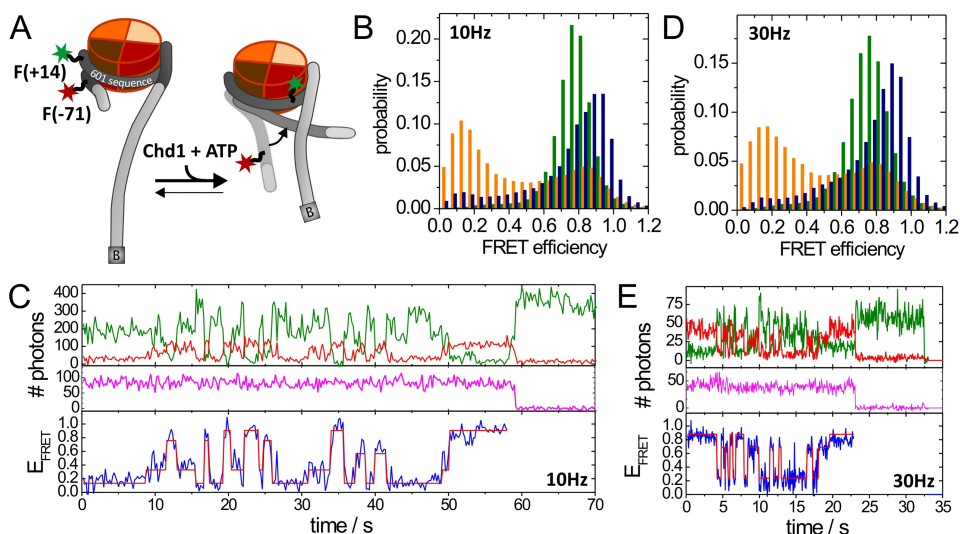


Figure 4.3.9.: Remodeling dynamics of F(-71)/F(+14) nucleosomes in the presence of 50 nM Chd1 and 150 μ M ATP as explored by smFRET at a time resolution of 200 ms (10 Hz) and 66 ms (30 Hz). The combined results of three and two independent measurements for 10 Hz and 30 Hz, respectively, are presented. (A) Cartoon illustrating the smFRET measurement of F(-71)/F(+14) nucleosomes before and after remodeling by Chd1. The distance between the two fluorophores was expected to increase during remodeling, however, even though an increase was observed intermediately, in the final state the distance was observed to be decreased (see panel B-C). A possible structural explanation is given. (B) Framewise FRET efficiency histograms of 10 Hz experiments are given for nucleosomes alone (green, 1898 molecules), nucleosomes showing a static FRET signal in the presence of Chd1 and ATP (blue, 1069 molecules) as well as nucleosomes showing dynamic transitions between different FRET states in the presence of Chd1 and ATP (orange, 409 molecules). Each histogram is built from unsmoothed time trajectories. (C) Exemplary dynamic FRET time trace of a single F(-71)/F(+14) nucleosome in the presence of Chd1 and ATP acquired at a time resolution of 200 ms. Time trace of donor fluorescence (green, upper panel) and acceptor fluorescence upon green excitation (red, upper panel) are shown together with the acceptor fluorescence upon direct excitation at 637 nm (magenta, middle panel). In the lower panel, the computed FRET efficiency (blue) is presented together with the time dependent transitioning between different FRET states as identified by HMM analysis (red). (D) Framewise FRET efficiency histograms of 30 Hz experiments are given for nucleosomes alone (green, 1219 molecules), nucleosomes showing a static FRET signal in the presence of Chd1 and ATP (blue, 637 molecules) as well as nucleosomes showing dynamic transitions between different FRET states in the presence of Chd1 and ATP (orange, 135 molecules). Each histogram is built from unsmoothed time trajectories. (E) Exemplary dynamic FRET time trace of a single F(-71)/F(+14) nucleosome in the presence of Chd1 and ATP acquired at a time resolution of 66 ms (color coding of traces as in panel C).

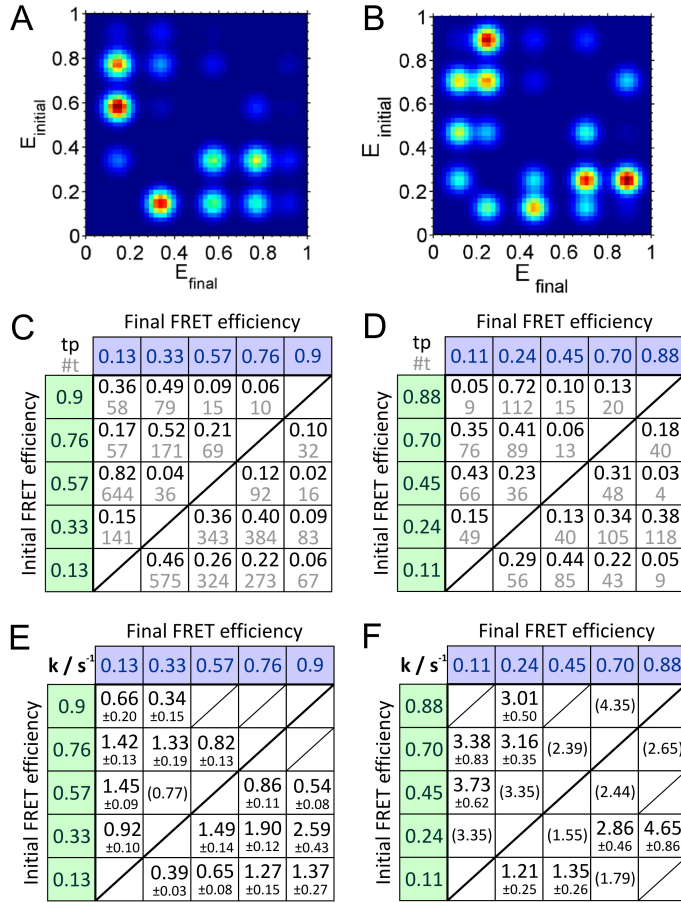


Figure 4.3.10: HMM analysis of dynamic FRET traces during remodeling of F(-71)/F(+14) nucleosomes with 50 nM Chd1 and 150 μ M ATP. Analysis of 10 Hz data is compared with the analysis of 30 Hz data. (A,B) TDPs for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from HMM analysis of a total of 409 dynamic traces showing 3907 transitions (10 Hz data, panel A) and of a total of 135 traces showing 1033 transitions (30 Hz data, panel B), respectively. A five-state hidden Markov process was assumed for both analysis. (C,D) Matrix presenting transition probabilities for all detected transitions (tp , black numbers) together with the absolute occurrence of each transition ($\#t$, gray numbers) for the 10 Hz experiments (panel C) and the 30 Hz experiments (panel D). Note that the sum of all transition probabilities originating from one state (i.e. all probabilities in one line of the matrix) has to be 1. (E,F) Matrices presenting transition rates and their standard deviations in s^{-1} for all detected transitions for the 10 Hz experiments (panel E) and the 30 Hz experiments (panel F). Each rate was extracted from a mono-exponential fit to the cumulative distribution of dwell times of each transition. The error is given as standard deviation of the rate from randomly chosen subsets of 30 % of molecules undergoing the same transition. Rates in parenthesis have only low statistics due to a low occurrence of the corresponding transition.

faster in the trajectory with 66 ms time resolution.³ In order to quantify the observed kinetics, global HMM analysis of dynamic time traces was performed (Figure 4.3.10). A five-state hidden Markov process was assumed, since with a lower number of states the initial FRET state at about 75 % could not be detected. Analysis of a total of 409 dynamic traces showing 3907 transitions (10 Hz data, Figure 4.3.10A) and of a total of 135 traces showing 1033 transitions (30 Hz data, Figure 4.3.10B) yielded similar TDPs for both time resolutions and transition probabilities were extracted (Figure 4.3.10 panel C, 10 Hz; panel D, 30 Hz). Five different states were detected at approximately 10%, 30%, 50%, 75 % and 90 % FRET and all possible transitions could be observed. The transition probability matrices for both time resolutions are partly consistent: The state at ~75 % FRET most likely transits to the state at ~30 % FRET and the state at ~50 % most likely transfers to the state at ~10 %. Moreover, from the lowest state at ~10 % FRET, transitions to the states at ~30 % and ~50 % are most strongly populated. For the 10 Hz data, transitions from the ~30 % FRET state back to ~75 % or ~50 % FRET were most probable. Taken together, this indicates a dynamic cycling between the four lower FRET states. One major difference between the HMM results for the two different time resolutions is the absolute occurrence of transitions into and from the FRET state at ~90 % of static, remodeled nucleosomes. At a time resolution of 10 Hz, these transitions were only rarely detected (see absolute occurrence in Figure 4.3.10C), whereas they were found to be most highly populated at a time resolution of 30 Hz (see absolute occurrence in Figure 4.3.10D). The fact that transitions to the final state of 90 % FRET were only rarely detected in the 10 Hz experiments (2-10 % probability) even though this state presents the final state of the overall remodeling process (Figure 4.3.9B, blue histogram) suggests that these transitions are irreversible on the time scale of 200 ms. However, the fact that transitions in and out of the ~90 % FRET state are frequently detected at a higher time resolution suggests that this state can also exist in an instable, short-lived form and is then frequently populated but not detectable at a lower time resolution. Such a situation could be explained if a conformational change in the nucleosomal DNA brings the two fluorophores at initial positions F(+14) and F(-71) close together, however the resulting conformation is instable in the presence of Chd1. Only upon dissociation of Chd1 away from the nucleosome, the conformation with a FRET efficiency of 90 % FRET is stable and therefore the transition to this state irreversible, because Chd1 is not present anymore to induce further conformational dynamics.

In addition to the transition probabilities, transition rates were extracted from the TDPs of experiments at both time resolutions and are presented in Figure 4.3.10 panel E (10 Hz) and panel F (30 Hz). While the rates ranged from 0.34 - 2.59 s⁻¹ in the 10 Hz experiments, 2-10 times higher rates ranging from 1.21 - 4.65 s⁻¹ were found for the measurements with a time resolution of 66 ms. This demonstrates that remodeling kinetics occur on a faster time scale than 200 ms and therefore an acquisition rate higher than 10 Hz would be

³Note that time trajectories are shown in the same scale and the 30 Hz trace is only half as long as the 10 Hz trace.

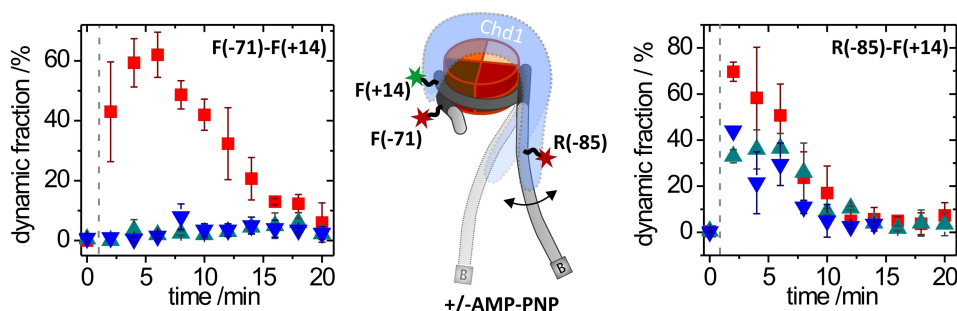


Figure 4.3.11.: Fraction of nucleosomes exhibiting FRET dynamics per movie (duration 120 s) presented for F(-71)/F(+14) nucleosomes (left panel) and for R(-85)/F(+14) nucleosomes. The data point at 0 min shows the fraction of dynamic nucleosomes in the absence of Chd1 and all following data points show the fraction of dynamic nucleosomes in the presence of 50 nM Chd1 and 150 μ M ATP (red squares), 50 nM Chd1 and 150 μ M AMP-PNP (cyan upwards triangles) and 50 nM Chd1 without any ATP analog (blue, downward triangles). Each data point represents the average of three independent measurements and the standard deviation is given by horizontal bars. The cartoon in the center illustrates the label positions and the observed dynamics with Chd1 and AMP-PNP or without ATP analog.

required in order to detect all transitions. As a consequence, the transitions detected at 200 ms time resolution presumably oftentimes do not follow mono-exponential decay kinetics and thus, the values of transition rates obtained from mono-exponential fits to the histogram of dwell times of the transitions need to be treated with care. However, even though the time resolution of 200 ms is too low to detect all remodeling dynamics, no markedly artificial FRET state due to averaging was created at the lower time resolution as visible by a comparison of the framewise FRET efficiency histograms of dynamic nucleosomes for the measurements at both time resolution (Figure 4.3.9B compared to D). The intermediate FRET state at 38 % was even slightly higher populated in the measurement with a higher time resolution, which suggests that more short-lived intermediate FRET states exist, that could not be detected at a time resolution of 200 ms.

Chd1 alters conformation of DNA linker on nucleosome entry site but not exit site

In a next set of experiments, F(-71)/F(+14) nucleosomes were exposed to 50 nM Chd1 in the presence or absence of 150 μ M AMP-PNP. However, unlike R(-85)/F(+14) nucleosomes, F(-71)/F(+14) nucleosomes did not show any FRET dynamics in the presence of Chd1 alone or with AMP-PNP and also the FRET efficiency histogram of static nucleosomes did not change compared to that of nucleosomes alone. This contrast is visualized in Figure 4.3.11, which summarizes the fraction of nucleosomes exhibiting FRET dynamics per movie (duration 120 s) over the course of a measurement for F(-71)/F(+14) nucleosomes (left panel) and for R(-85)/F(+14) nucleosomes (right panel).

While upon addition of 50 nM Chd1 and 150 μ M ATP (red squares) F(-71)/F(+14) nucleosomes showed a drastic increase in the dynamic fraction from about 0 % in the

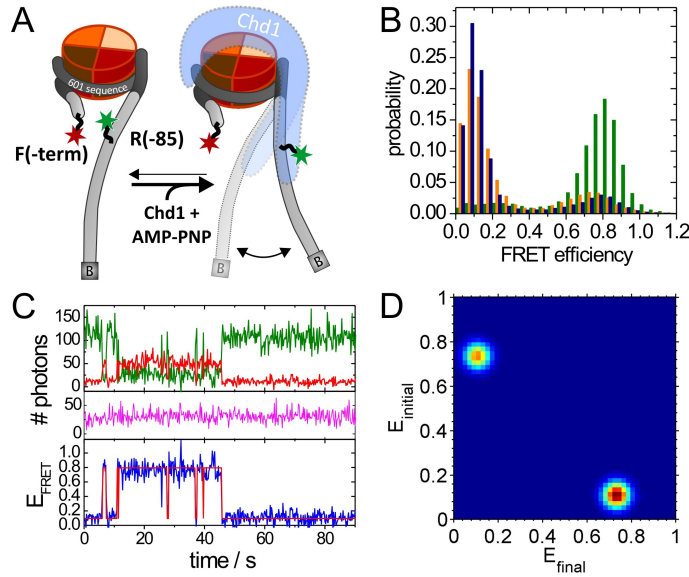


Figure 4.3.12: Dynamics of F(-term)/R(-85) nucleosomes in the presence of 50 nM Chd1 and 150 μ M AMP-PNP explored by smFRET. The combined results of two independent measurements are presented. (A) Cartoon illustrating F(-term)/R(-85) nucleosomes in the absence and in the presence of Chd1 and AMP-PNP. Observed dynamics of the long linker DNA (see panel C-D) in the presence of Chd1 are indicated. (B) Framework FRET efficiency histogram of nucleosomes alone (green, 417 molecules), static nucleosomes in the presence of Chd1 and AMP-PNP (blue, 673 molecules) as well as nucleosomes showing FRET dynamics in the presence of Chd1 and AMP-PNP (orange, 81 molecules). All histograms contain unsmoothed time trajectories. (C) Example of directly observed dynamic transitions between a high FRET and a low FRET state of a single F(-term)/R(-85) nucleosome induced by Chd1 and AMP-PNP. Time trace of donor fluorescence (green, upper panel) and acceptor fluorescence upon green excitation (red, upper panel) are shown together with the acceptor fluorescence upon direct excitation at 637 nm (magenta, middle panel). In the lower panel, the computed FRET efficiency (blue) is presented together with the time dependent transitioning between different FRET states as identified by HMM analysis (red). (D) TDP for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from HMM analysis of a total of 81 dynamic traces showing 289 transitions.

absence of Chd1 and ATP (0 min) to about 60 %, no dynamics were observed in the presence of Chd1 without ATP analog (blue downward triangles) or with 150 μ M AMP-PNP (cyan upwards triangles). In contrast, R(-85)/F(+14) nucleosomes showed an increase in dynamic fraction from about 0 % (0 min) to about 70 % not only upon addition of Chd1 and ATP, but also in the presence of Chd1 alone or Chd1 and AMP-PNP a reasonable increase of the dynamic fraction to about 40 % could be detected. This shows that Chd1 alone or with AMP-PNP induces dynamic conformational changes in the long linker DNA, to which it is assumed to bind via its DNA-binding domain. In contrast, the conformation of the short linker DNA on the nucleosome exit site is not disturbed by Chd1 in the absence of ATP. In all cases where dynamic FRET was observed in the presence of Chd1, the dynamic fraction decreased back to 0 % over the course of the ex-

periment (12-20 min). For experiments with Chd1 and ATP, this decrease in the dynamic fraction could be explained such that dynamics in FRET are a sign for active remodeling, which is completed after a certain time. Yet, the fact that this decrease in the fraction of dynamic nucleosomes was observed also for R(-85)/F(+14) nucleosomes under conditions that prevent the remodeling reaction from being completed, such as with Chd1 alone or with AMP-PNP, shows that this explanation is not sufficient. I assume that the decrease in fraction of dynamic nucleosomes results from a loss of Chd1 activity when incubated at room temperature. Such a loss of activity might be the effect of reduced Chd1 binding due to precipitation from solution. This assumption is supported by the observation, that some FRET dynamics can be reestablished even after 20-40 min by supplying fresh remodeler to the immobilized nucleosomes (data not shown and [257]).

Finally, the conformation of both DNA linkers relative to each other upon binding of Chd1 was examined. Therefore, smFRET experiments were performed with F(-term)/R(-85) nucleosomes in the presence of 50 nM Chd1 and 150 μ M AMP-PNP (Figure 4.3.12). The FRET efficiency histogram of nucleosomes alone showed a single peak at a high FRET efficiency of 81 % indicating a relatively closed structure of the nucleosome with a short distance between the two fluorophores attached to extranucleosomal DNA on opposite sides of the nucleosome (Figure 4.3.12B, green histogram). In the presence of Chd1 and AMP-PNP, the FRET efficiency histogram of static nucleosomes drastically changed in that the peak at 80 % FRET became a small side peak with a probability of 18 %, whereas a major peak was visible at a low FRET efficiency of 9 % (82 % probability) (Figure 4.3.12B, blue histogram). Besides, 12 % of nucleosomes showed dynamics in the presence of Chd1 and AMP-PNP and the resulting FRET efficiency distribution was similar to the distribution of static nucleosomes showing a major peak at 9 % FRET and a side peak at 80 % FRET (Figure 4.3.12B, orange histogram). Dynamic smFRET trajectories and the corresponding TDP resulting from HMM analysis of 81 dynamic traces showing 289 transitions revealed dynamic switching between the two FRET states at 10 % and 80 % FRET. Hence, Chd1 induces a conformational change in the linker DNA such that a more open structure is adopted by the nucleosome with both linkers further apart. Moreover, nucleosomes are dynamically switching between the closed and open conformation, consistent with the results of previous experiments with R(-85)/F(+14) nucleosomes in the presence of 50 nM Chd1 and 150 μ M AMP-PNP. From the TDP a cumulative distribution of dwell times was extracted for each transition. However, only the transition from 80 % \rightarrow 10 % FRET could be well described by a mono-exponential fit yielding a transition rate of $k_{0.8 \rightarrow 0.1} = 0.38 \text{ s}^{-1}$. The cumulative distribution of dwell times for the reverse transition was best described by a bi-exponential decay function ($k_{0.1 \rightarrow 0.8,1} = 4.00 \text{ s}^{-1}$ and $k_{0.1 \rightarrow 0.8,2} = 0.09 \text{ s}^{-1}$) suggesting that two different pathways are underlying the transitioning of nucleosomes from the low FRET state at 10 % back to the initial high FRET state at 80 %.

4.3.2. Discussion

I developed a smFRET assay using mono-nucleosomal constructs labeled with FRET pairs at different sites of the nucleosomal DNA to follow repositioning of individual nucleosomes by Chd1 in real-time. In combination with Hidden Markov Modeling analysis of dynamic FRET trajectories, this assay allowed me to quantitatively characterize structural dynamics of the nucleosomal DNA during the remodeling process. In the presence of ATP, Chd1 translocates nucleosomal DNA in a bidirectional and stepwise manner on a time scale of hundreds of milliseconds to seconds. The dynamics were independent of Chd1 concentration in the investigated concentration range and observed even with prebound Chd1 in the absence of remodeler in solution, demonstrating that continuously bound Chd1 processively moved DNA. Moreover, even in the absence of ATP, I observed Chd1 induced conformational changes in the DNA, which I could attribute to the long linker DNA on the nucleosome entry site. In the following, I want to combine the smFRET results for the different nucleosomal constructs and try to structurally interpret the observed different FRET states with respect to the underlying intermediate nucleosome structures. By doing so, I will present one possible model for the mechanism of nucleosome remodeling by Chd1.

In order to develop a possible model for the mechanism of Chd1 remodeling, a number of considerations were made:

Two nucleosomal constructs were examined during Chd1 catalyzed remodeling, of which one was sensitive to dynamics in the nucleosomal DNA at the nucleosome entry site (R(-85)/F(+14)) and the other one was able to monitor dynamics of the nucleosomal DNA at the nucleosome exit site (F(-71)/F(+14)). Several differences between the two nucleosomal constructs could be detected: While R(-85)/F(+14) nucleosomes evolved through 4 different FRET states, F(-71)/F(+14) nucleosomes captured 5 different states during remodeling by Chd1. Moreover, all transitions between the five FRET states of F(-71)/F(+14) nucleosomes required the energy of ATP hydrolysis, whereas R(-85)/F(+14) nucleosomes switched between two of four states even in the absence of ATP. Finally, static, remodeled R(-85)/F(+14) nucleosomes showed a FRET histogram with one distinct peak centered at 86 % FRET (Figure 4.3.4B), whereas remodeled F(-71)/F(+14) nucleosomes showed a broad distribution with a peak at 90 % FRET and a broad shoulder comprising 38 % of the data at lower FRET efficiencies (Figure 4.3.9B). A proposed remodeling mechanism has to account for these discrepancies between the two differently labeled nucleosomal constructs.

Moreover, in this study all fluorescent labels were attached to the nucleosomal DNA, which is translocated over the histone octamer surface during remodeling by Chd1. Therefore, both labels of a FRET pair are expected to move at a given time during the remodeling process. If remodeling proceeded via the proposed twist-diffusion model [107, 161], in which the DNA propagates over the octamer surface in a corkscrew manner (see Figure 1.2.5B), a continuous change in FRET efficiency would be expected, if the corkscrew

movement was not substantially faster than the measurement time resolution. The fact that for both nucleosomal constructs, R(-85)/F(+14) as well as F(-71)/F(+14), the observed FRET efficiency changed in a stepwise manner supports the idea that remodeling by Chd1 involves formation of a DNA loop and its propagation over the histone octamer surface. Hence, I assumed a loop-recapture model for the remodeling mechanism of Chd1.

Previous work showed that nucleosomes are preferentially repositioned by a multiple of 10 bp from their initial position, consistent with the 10 bp periodicity of DNA-histone contacts within the nucleosome [166, 165]. Moreover, in a recent smFRET study exploring remodeling of mono-nucleosomes by the ISWI-type remodeler ACF [14], intermediate remodeling steps of 3 and 7 bp were reported, however nucleosomal states were stable only when translocation by a multiple of 10 bp had occurred. Considering this knowledge from previous studies and the conclusions from the data presented here, I acted on the assumption that continuously bound, monomeric Chd1 repositions nucleosomes by a loop-recapture mechanism with a step size of 10 bp. Moreover, I assumed preferential translocation of the initial 6-N-47 nucleosomes by a total of 20 bp to yield 26-N-27 nucleosomes, consistent with a previous footprinting study, where Chd1 remodeling of end-positioned nucleosomes (0-N-54) was found to yield 26-N-28 nucleosomes [128].

Kinetic rates and probabilities of the individual transitions extracted from dynamic smFRET traces by global HMM analysis were not taken into account for developing a model of the Chd1 remodeling mechanism, since they were not reliable enough due to several reasons. Most experiments were performed at a time resolution of 200 ms, which was shown to be too low in order to detect all remodeling dynamics (see Figure 4.3.10 and corresponding discussion). As a consequence, transitions cannot be assumed to follow mono-exponential decay kinetics since multiple processes might be underlying a single detected transition. Moreover, the examined system was observed not to be in equilibrium, since even in case of the two-state system of R(-85)/F(+14) nucleosomes in the presence of Chd1 without ATP the forward and reverse transitions were not populated equally (see Figure 4.3.6). This might be the effect of an insufficiently long observation time due to photobleaching of the fluorophores.

In the following I will describe a speculative model for nucleosome remodeling by Chd1, which is able to explain the presented smFRET data (Figure 4.3.13). However, one has to keep in mind that the presented model is only one possible interpretation of the data and future experiments investigating Chd1 induced remodeling, especially at a higher time resolution than presented here, should aim for testing this model and either confirming and refining or even disproving it.

In its initial state, the nucleosome is mostly static with a conformation resembling the

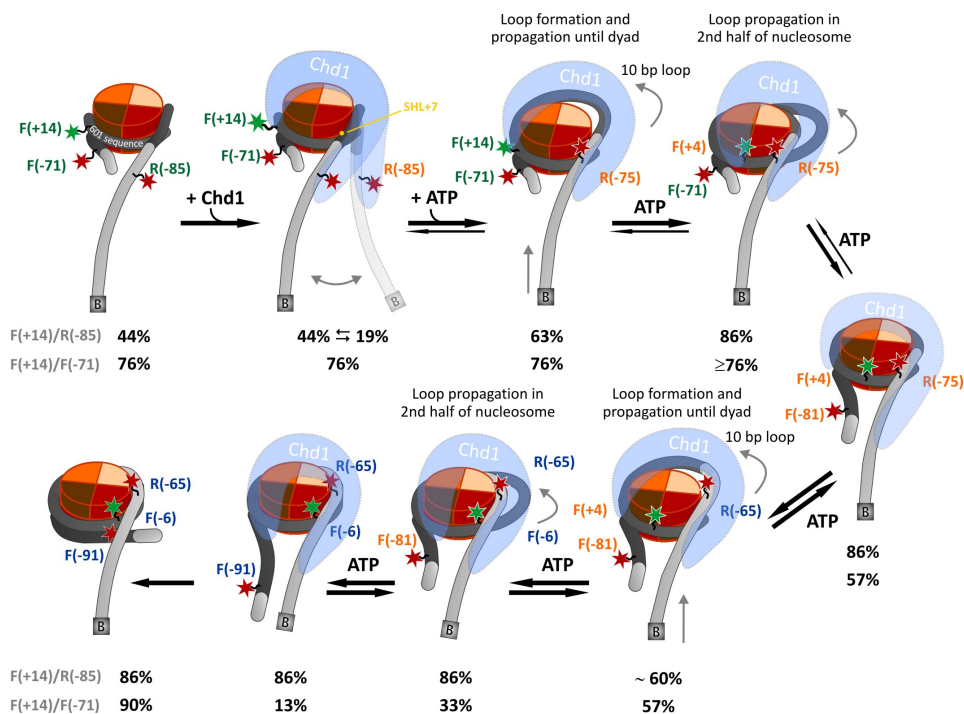


Figure 4.3.13.: Schematic illustration of the mechanism, by which Chd1 remodels end-positioned mono-nucleosomes towards the center of the DNA as inferred from smFRET dynamics data. A loop-recapture model is assumed with a total repositioning by 20 bp in two 10 bp steps. Chd1 is sketched in blue. Fluorescent dyes are indicated by stars (green, Tamra; red, Alexa647) and their position is labeled in green when at its initial position, in orange when moved to an intermediate position and in blue when moved to the final position. FRET efficiencies of R(-85)/F(+14) and F(-71)/F(+14) nucleosomes are shown below the nucleosome cartoons. The first five cartoons make up the first 10 bp remodeling step and the last four cartoons represent the second. All transitions but the first one are assumed to require ATP. Except for the last step, all steps are assumed to be able to proceed in both directions. During the last transition, Chd1 dissociates from the nucleosome and hence, this transition is irreversible. See main text for a detailed description of the remodeling mechanism.

structure that was captured by x-ray crystallography [5, 256]. Upon addition of Chd1 and binding to the nucleosome, Chd1 induces a dynamic conformational change in the long linker DNA on the nucleosome entry site, presumably via its DNA-binding domain. This conformational change does not require ATP and involves movement of the linker DNA away from the nucleosome core as observed by the dynamic switching of R(-85)/F(+14) nucleosomes between the initial 44 % FRET state and the state at a lower FRET efficiency of 19 %. Also the ISWI remodeler ACF was previously shown to perturb histone-DNA contacts near the nucleosome entry site during binding to the nucleosome in the absence of ATP [170]. In contrast, the nucleosome exit site probed by the F(-71)/F(+14) dye pair remains static. Chd1 then uses the energy of ATP hydrolysis to break the first histone-

DNA contacts at the entry site of the nucleosome and reestablish an interaction at the outermost contact site (superhelical location +7) between histones and a position initially 10 bp outside of the nucleosome core ($\sim R(-83)$). In this way, a 10 bp DNA loop is created on the histone octamer surface and the dye at position $R(-85)$ moves to position $R(-75)$ and thus closer to $F(+14)$, observable by an increase in FRET efficiency from 19/44 % to 63 % FRET. Yet, the distance between positions $F(-71)$ and $F(+14)$ remains unchanged. In the next step, the DNA loop propagates like a wave over the octamer surface, thereby breaking histone-DNA contacts at the leading edge of the loop and replacing them at the lagging edge of the loop. In this way, only one or two histone-DNA contacts have to be disrupted at a time, which renders the remodeling process energy-efficient. During loop propagation, $F(+14)$ and $F(-71)$ are consecutively translocated to position $F(+4)$ and $F(-81)$, respectively, which is observed as an increase in FRET efficiency of $R(-85)/F(+14)$ nucleosomes from 63 to 86 % followed by a decrease in FRET efficiency of $F(-71)/F(+14)$ nucleosomes from 76 to 57 %. The initial movement of $F(+14)$ to $F(+4)$ is not assumed to lead to a transition of $F(-71)/F(+14)$ nucleosomes into a different state than that at 76 % FRET, since the overall distance between the two dyes only marginally changes. Only a very transient increase in FRET efficiency might be observed. With this, the first 10 bp remodeling step is finished and $R(-85)/F(+14)$ nucleosomes, now $R(-75)/F(+4)$, have reached a state at 86 % FRET, whereas $F(-71)/F(+14)$ nucleosomes, now $F(-81)/F(+4)$, have transited into a state at 57 % FRET.

Subsequently, the whole process of sequential loop formation and propagation is repeated in a second 10 bp remodeling step to yield an overall translocation of nucleosomes by 20 bp and the final dye pair positions $R(-65)/F(-6)$ and $F(-91)/F(-6)$. Interestingly, beside a transient decrease in FRET efficiency during loop propagation, $R(-75)/F(+4)$ nucleosomes do not undergo another overall FRET efficiency change during the transition to $R(-65)/F(-6)$, since the dyes are translocated parallelly within the nucleosome core. However, the FRET efficiency of $F(-81)/F(+4)$ nucleosomes further decreases from 57 % FRET via 33 % FRET (loop propagation past the first gyre) to 13 % FRET (loop release at nucleosome exit site) during the transition to $F(-91)/F(-6)$. After this second 10 bp translocation of the nucleosome, a center-positioned nucleosome is created with 27 bp and 26 bp long DNA linkers on the entry and exit site, respectively. The second 10 bp remodeling step is assumed to occur with a relatively high probability also in the reverse direction because the nucleosome has sufficiently long DNA linkers on both sides so that the two possible binding orientations of Chd1 to the nucleosome are assumed to be equally probable.

The following increase in FRET efficiency of $F(-91)/F(-6)$ nucleosomes from 13 % to 90 % FRET is assumed to be due to closing of the linker DNA on the exit site towards the nucleosome entry site upon dissociation of Chd1 away from the nucleosome. However, the closing is only a speculation, which remains to be verified by smFRET investigation of initially center-positioned $F(-91)/F(-6)$ nucleosomes in the absence and presence of Chd1.

4 Single molecule FRET studies of nucleosome remodeling by Chd1

Further remodeling by a third 10 bp step (not shown) can follow yielding 36-N-17 nucleosomes, however it is not very likely since the favored center position of the nucleosome would be abandoned and since Chd1 was shown to be unable to move nucleosomes to positions closer than about 15 bp from a DNA end [128].

Each step in the proposed remodeling mechanism (except for the last step involving dissociation of Chd1 from the nucleosome) is assumed to be able to proceed in either direction, since dynamic switching in between the neighboring FRET states was observed. Moreover, all conformational changes except for the initial linker DNA dynamics are assumed to require the energy of ATP hydrolysis, since a dependency of the corresponding transition rates on the concentration of ATP was found. The fact that in spite of the bidirectionality of each step a relatively well defined remodeling product (visible as a well defined FRET efficiency) with a center-positioned nucleosome is produced, is attributed to the transition rates being faster for movement towards the center of the DNA and slower for movement away from the center (Figure A.2.4 and Figure 4.3.10). This discrepancy in the magnitude of rates can be attributed to the previously reported affinity differences of Chd1 in binding to extranucleosomal DNA [164, 128, 258, 152]. Chd1 was shown to sense the presence of extranucleosomal DNA through its DNA-binding domain and preferentially bind to the longer linker DNA of a nucleosome. As a result, Chd1 translocates the nucleosome towards this longer stretch of linker DNA thereby reducing its length and creating a more centered nucleosome position. Since the binding affinity of Chd1 is higher for nucleosomes with a longer extranucleosomal DNA, the remodeler is expected to dissociate most probably from a remodeled nucleosome with shorter linker DNA. In the smFRET experiments, I observed dynamic bidirectional movement of the DNA carried out by a continuously bound Chd1 enzyme. Therefore, I envision that a single Chd1 remodeler anchored to the nucleosome core presumably via its ATPase and chromodomain is able to sense the linker DNA at both sides of a nucleosome. It preferentially binds to the longer DNA linker and translocates the nucleosome towards it. For bidirectional translocation, also the ATPase domain might have to flip its orientation. In this way, a dynamic equilibrium between different nucleosome positions is created, however center positions are most highly populated. A similar mechanism has been proposed earlier based on biochemical nucleosome binding and sliding assays [128].

The proposed model for the Chd1 remodeling mechanism is able to explain the differences between the FRET efficiency distributions of remodeled R(-85)/F(+14) and F(-71)/F(+14) nucleosomes: Remodeling products of 10 bp and 20 bp remodeling are thought to exist in dynamic equilibrium. In case of remodeling by 10 bp, R(-85)/F(+14) nucleosomes have already reached the final FRET efficiency of 86 %, whereas F(-71)/F(+14) nucleosomes are in an intermediate FRET state at 57 % FRET. This explains why the FRET efficiency distribution of remodeled R(-85)/F(+14) nucleosomes showed one dis-

tinct peak centered at 86 % FRET (Figure 4.3.4B), whereas the histogram of remodeled F(-71)/F(+14) nucleosomes showed beside the peak at 90 % FRET a broad shoulder comprising 38 % of the data at lower FRET efficiencies (Figure 4.3.9B). Moreover, the model is able to explain the different numbers of states for the two nucleosomal constructs.

On the basis of the presented smFRET data, it cannot be ruled out that DNA bulges larger than 10 bp are formed during Chd1 remodeling instead of the 10 bp loops assumed in the presented model. Formation and propagation of a bigger DNA loop would result in a larger remodeling step size and could be observed by switching between FRET states that are further apart on the FRET scale. Thus, one speculative explanation for the observed switching between non-neighboring FRET states is the remodeling via an alternative mechanism that involves larger (e.g. 20 bp) DNA bulges.

Another explanation for the observation of transitions between non-neighboring states could be that some transitions between neighboring states were faster than the applied measurement time resolution of 200 ms and were therefore observed as a direct transition between non-neighboring states with longer dwell times. However, the fact that even in smFRET experiments with a three times faster time resolution transitioning between non-neighboring states and a similar relative occupancy of intermediate and extreme FRET states was observed argues against this explanation.

A similar remodeling mechanism as the one presented here for Chd1 involving dynamic equilibrium between different nucleosome positions and faster rates for movement of nucleosomes towards the center of DNA was proposed previously for the ISWI type remodeler ACF [259]. Just like Chd1, ACF translocates mono-nucleosomes towards the center of the DNA and modifies nucleosomal arrays by equally spacing the nucleosomes. ACF was shown to act as a dimer with two monomers bound on the two opposite sides of a nucleosome and oriented for translocation in opposing directions [172, 14]. Coordinated action of the two monomers was thought to allow for processive back-and-forth movement of the nucleosome. However, bidirectional movement by ACF was observed only for mono-nucleosomal templates containing 303 bp DNA (78-N-78), whereas mono-nucleosomes containing 228 bp DNA (3-N-78) were remodeled in a unidirectional manner towards the center of the DNA [14]. In contrast, Chd1 is able to bidirectionally move mono-nucleosomes already on a 200 bp long DNA template (47-N-6) (4.3.1) suggesting that the DNA binding domain of ACF requires a longer stretch of extranucleosomal DNA than that of Chd1. This might be due to the fact that the physical dimension of the DNA-binding domain of ACF is larger compared to the DNA-binding domain of Chd1, which lacks the HAND domain and has a shorter spacer helix [150]. Moreover, in contrast to the single subunit remodeler Chd1, ACF is a heterodimeric protein consisting of the catalytic subunit Snf2h and the attendant subunit Acf1, which is supposed to influence remodeling. The distinct requirements for extranucleosomal DNA of ACF and Chd1 might explain the previous observation that ACF and Chd1 assemble chromatin with distinct

nucleosome repeat lengths [127]: while Chd1 was found to introduce nucleosome spacing with a repeat length of 162 bp, ACF spaced nucleosomes with a repeat length of 175 bp. Another potential difference between Chd1 and ACF that might contribute to the observed discrepancy with respect to remodeling of mono-nucleosomes with different linker DNA lengths is the monomeric versus dimeric binding behavior. Even though multiple Chd1 remodelers are able to bind a single mono-nucleosome, as visible by a supershift of nucleosome bands in a Chd1 binding assay (Figure 4.6.6 and [149]), Chd1 is assumed to bind mono-nucleosomes predominantly as a monomer. Evidence therefor stems from a smFRET based Chd1 counting assay and statistical analysis, in which the presence of mainly one Chd1 remodeling enzyme per nucleosome was detected (see section 4.4.1). In contrast, ACF was shown to bind nucleosomes as a dimer [172, 14]. Hence, bidirectional movement of the nucleosome in the case of ACF seems to result from the coordinated action of two remodelers bound to the opposite sites of a nucleosome, whereas in the case of Chd1 it stems from a single, continuously bound remodeler that is able to sense both DNA linkers and move the nucleosome in both directions. It remains to be shown whether bidirectional movement by a single Chd1 remodeler involves flipping of the orientation of the DNA-binding domain or even of the complete remodeler relative to the nucleosome or a change in translocation directionality of the ATPase domain.

While the nucleosomal FRET constructs used by Blosser et al. [14] allowed for the direct observation of nucleosome translocation by a decrease in FRET efficiency, they were not sensitive to structural changes within the nucleosomal DNA. Therefore, the origin of the ATP-dependent kinetic pauses observed during translocation by ACF remained unclear. I used nucleosomal FRET constructs that allowed for the monitoring of dynamics within the nucleosomal DNA and I observed a gradual movement of nucleosomal DNA at the entry site, around the dyad and at the exit site of the nucleosome. This directly suggests that remodeling by Chd1 proceeds via formation and propagation of a DNA loop.

4.4. Architecture of a Chd1-nucleosome complex

4.4.1. Results

Assembly of functional, labeled Chd1-nucleosome complex suited for NPS investigation

To probe the architecture of Chd1-nucleosome complexes by global NPS analysis, I assembled complexes from donor labeled, end-positioned mono-nucleosomes and acceptor labeled truncated Chd1 1-1305 Δ C (referred to as Chd1 in the following) in the presence of AMP-PNP. I attached assembled complexes to the surface of microfluidic chambers and investigated them by smFRET using total internal reflection fluorescence microscopy. Mono-nucleosomes contained the nucleosome positioning sequence 601 [111] to obtain a strongly positioned nucleosome core, as well as DNA segments of 47 bp and 6 bp extending from the boundaries of the nucleosome core at the entry and exit site, respectively (6-N-47). Moreover, a biotin label was attached to the end of the long extranucleosomal DNA for immobilization in the microfluidic smFRET measurement chamber. Nucleosomes were labeled with Tamra as satellite dye molecule on one of the following positions of the nucleosomal DNA: F(+14), F(+2), F(-15), F(-64), R(-60) or R(-66) (Figure 4.4.1A and Figure 4.6.1). Single-cystein mutants of Chd1 were labeled with Alexa647 as antenna dye molecule at one of two positions in the tandem chromodomains (C190, C293), at one of three positions in the Snf2 ATPase domain (C365, C647, C732) or at one of three positions in the DNA-binding domain within the HL1 domain (C1091), the spacer helix (C1149) or the SLIDE domain (C1159) (Figure 4.4.1B).

I verified that introduction of fluorescent labels at different sites of Chd1 and the nucleosomal DNA did not impair Chd1 remodeling activity by performing a remodeling assay using Tamra labeled mono-nucleosomes and Alexa647 labeled Chd1 (section 4.6.12, exemplary data shown in Figure 4.6.7). Fluorescently labeled, mutant Chd1 successfully repositioned labeled mono-nucleosomes from the end towards the center of the DNA, as visible by an upshift of the initial nucleosome band in native gel analysis. Furthermore, I demonstrated successful assembly of double-labeled Chd1-nucleosome complexes by using a Chd1 binding assay (section 4.6.11, exemplary data shown in Figure 4.6.6). The influence of different non-hydrolyzable ATP analogs on the nucleosome binding activity of Chd1 was probed and since, in comparison to ATP- γ S, ADP-BeF_x or no ATP analog, AMP-PNP most strongly promoted binding, it was used in all smFRET experiments (section 4.6.11). Since AMP-PNP cannot be hydrolyzed but can bind to the ATP-binding pocket of the Chd1 ATPase domain, stalled remodeler-nucleosome complexes are formed.

Chd1-nucleosome complex contains predominantly a single Chd1 remodeler

Chd1 binding assays indicated that more than one Chd1 molecule can interact with one nucleosome if Chd1 is supplied in excess (see supershifts in Figure 4.6.6). This observation

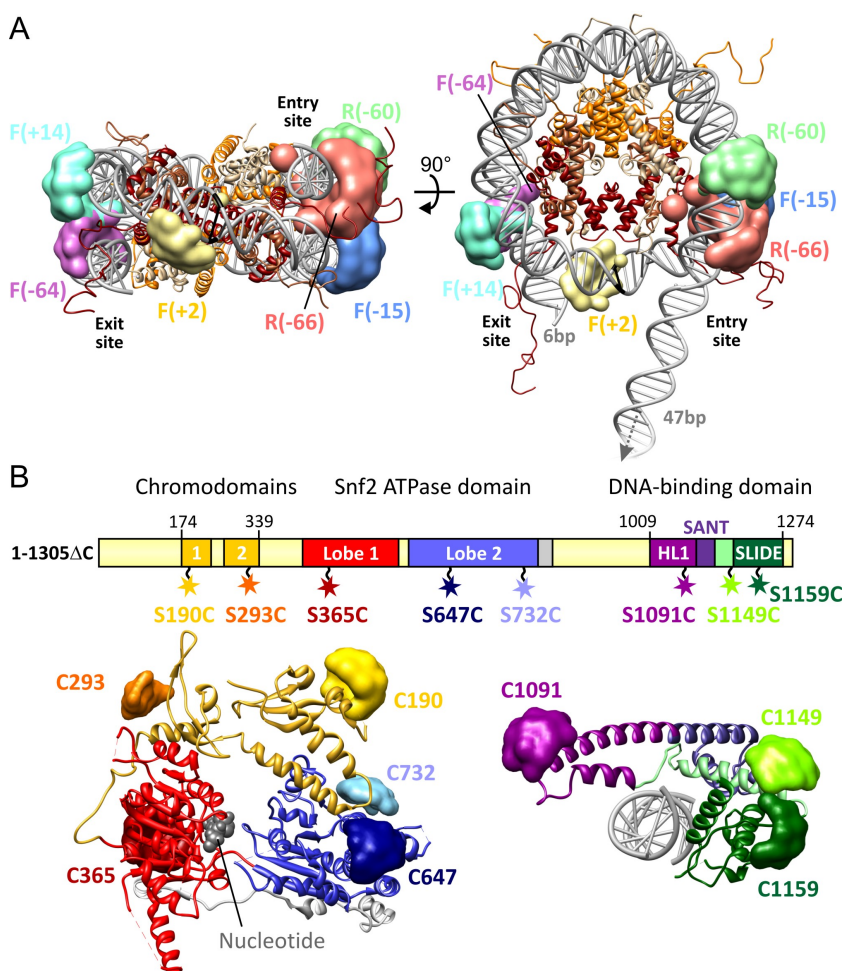


Figure 4.4.1.: Experimental design of labeled Chd1-nucleosome complex suited for global NPS analysis. (A) Position priors of satellite dye molecules attached to the nucleosomal DNA are shown in the reference coordinate system of the nucleosome core particle (pdb ID: 3LZ0 (structure of nucleosome containing 601 DNA [256]; only DNA of this structure is shown here), 1KX5 (nucleosome structure with visible histone tails [99]; only histone octamer of this structure is shown here, which was perfectly aligned with histone octamer in 3LZ0 structure); H3, red, H4, brown, H2A, tan, H2B, orange). Linker DNA at nucleosome entry site (47 bp) and exit site (6 bp) are indicated. The color coding introduced for the different satellites will be used throughout this chapter. (B) Top: Domain structure of truncated Chd1 1-1305 Δ C is shown together with sites subjected to single cysteine point mutation for attachment of antenna dye molecules. Color coding introduced for different Chd1 domains and antennas will be used throughout this chapter. Bottom left: For illustration purposes of antennas attached to the tandem chromodomains and to the ATPase domain, their accessible volumes were modeled and are shown in the respective crystal structure (pdb ID: 3MWY, [147]). The structures of both domains separately and relative to each other are expected to change when bound to a nucleosome (section 1.2.5). Bottom right: Position priors of antenna dye molecules on the DNA-binding domain are shown in their reference frame (pdb ID: 3TED, [150]), which is being located and oriented as a rigid domain relative to the nucleosome coordinate system during NPS docking analysis. All images were prepared in Chimera [151].

persuaded me to examine the stoichiometry of functional Chd1-nucleosome complexes. To this aim, I assembled Chd1-nucleosome complexes in microfluidic smFRET measurement chambers (section 2.2.2) using Alexa647 labeled Chd1-C365 (labeling efficiency 65 %) and F(+2)-Tamra nucleosomes and performed smFRET experiments in ALEX mode. First, donor labeled mono-nucleosomes were immobilized via their biotin anchor on the surface of microfluidic chambers followed by supplying acceptor labeled Chd1 to the chamber, incubating for some time and finally removing excess, unbound remodeler from solution by thorough rinsing. I used 50 nM Chd1 since this concentration yielded complete remodeling in the smFRET remodeling experiments presented above (section 4.3). The resulting fluorescence time trajectories were analyzed by counting the number of bleaching steps of acceptor Alexa647 per Chd1-nucleosome complex. Only those complexes that showed a single bleaching step in the donor channel were included in the analysis in order to avoid aggregates containing more than one nucleosomal DNA. Considering the labeling efficiency of 65 % for Chd1-C365, 43 % of all complexes would be expected to show two bleaching steps, if Chd1 consistently bound the nucleosome as a dimer. However, analysis of two independent experiments yielded in average 77 ± 9 % of complexes with one bleaching step, 14 ± 6 % with two and 9 ± 6 % of complexes with three or more bleaching steps. In addition, not all Chd1-nucleosome complexes holding both acceptor and donor label exhibited FRET, however a FRET signal was believed to be a sign for properly assembled complexes. The fraction of complexes exhibiting FRET, was much higher for those with one Alexa647 bleaching step (0.89 ± 0.01) than those with two or more bleaching steps (0.11 ± 0.01) indicating that Chd1 multimers are mainly bound in an unfunctional way. Hence, this analysis revealed that functional Chd1-nucleosome complexes predominantly contained a single Chd1 remodeler. It cannot be excluded though that occasionally two remodelers interact with a nucleosome also in a functionally active manner.

Conformation of nucleosome core particle within Chd1-nucleosome complex

In order to use fluorophores attached to the nucleosomal DNA as satellites, their position had to be known within the laboratory coordinate system of the nucleosome core particle. I assumed nucleosomal DNA within the nucleosome core particle (strong positioning sequence 601) in Chd1-nucleosome complexes to be unchanged compared to its conformation in the crystal structure (pdb ID: 3LZ0, [256]). This assumption was tested by comparison of distance measurements between sites on the nucleosomal DNA in the absence and presence of Chd1 and AMP-PNP. The conformation of nucleosomal DNA at the nucleosome exit site was probed by smFRET measurements between F(-71)-Alexa647 and satellite F(+14)-Tamra and the conformation of the nucleosomal DNA at the entry site was tested by smFRET measurements between the outermost satellite on the entry site, R(-66)-Alexa647, and satellite F(+2)-Tamra close to the dyad as well as between satellites R(-60)-Alexa647 and F(-15)-Tamra (Figure 4.4.2). The conformation of nucle-

4 Single molecule FRET studies of nucleosome remodeling by Chd1

osomal DNA at the nucleosome exit site (Figure 4.4.2B) did not alter upon binding of Chd1 since the smFRET histograms in the absence and in the presence of Chd1 were similar with a maximum FRET efficiency of 75 % corresponding to a distance of ~ 50 Å ($R_0^{iso} = 60$ Å). The maximal distance between the accessible volumes of the dye attached to F(-71) and satellite F(+14) is 70 Å and the distance between the attachment points measured in the crystal structure is 37 Å. Hence, the distance determined by smFRET is conform with the crystal structure of the nucleosome core particle. However, the unimodal smFRET histogram for the entry site measurement between R(-66) and F(+2) (Figure 4.4.2C) shifted from a maximal FRET efficiency of 52 % in the absence of Chd1 to 35 % in the presence of remodeler and AMP-PNP. With R_0^{iso} of 60 Å, this FRET change corresponds to an increase in distance by ~ 7 Å from ~ 59 Å to ~ 66 Å upon Chd1 binding. In comparison, the maximal distance between the accessible volumes of satellites R(-66) and F(+2) is 75 Å and the distance between the attachment points measured in the crystal structure of the nucleosome containing the 601 sequence [256] is 47 Å. Hence, the measured inter-dye distance of 66 Å in the presence of Chd1 is still consistent with the crystal structure and the change in FRET in the presence of Chd1 could result from dislocation of at least one of the dye molecules relative to the fixed attachment sites due to interaction with Chd1. However, the possibility of a conformational change in the nucleosomal DNA close to the entry site (e.g. disruption of the first histone-DNA contact at position R(-65)) due to binding of Chd1 cannot be excluded. Notably, the width of the histograms for the measurements between R(-66) and F(+2) with and without Chd1 were considerably larger than those for the histograms of measurements between F(-71) and F(+14), indicating that the outer nucleosomal DNA at the entry site of the nucleosome shows an increased flexibility. Finally, the FRET efficiency histogram for the measurement between R(-60) and F(-15) did not change in the presence compared to the absence of Chd1 and was narrow with a maximum at 81 % FRET in both cases (Figure 4.4.2D), showing that the position of R(-60) is not altered in the presence of Chd1. The FRET efficiency of 81 % corresponds to a distance of 47 Å ($R_0^{iso} = 60$ Å), which is consistent with the maximal distance of the accessible volumes of both satellites of 63 Å. Hence, if nucleosomal DNA at the nucleosome entry site is displaced in the presence of Chd1, at maximum the first histone-DNA contact can be broken and only about 10 base pairs can be involved in the conformational change. Therefore, I nevertheless used the nucleosome core particle crystal structure as a model for the location of satellites attached to the nucleosomal DNA. Yet, the position prior for satellite R-66 was enlarged by 7 Å (complete length of flexible linker: 20 Å) to account for a possible displacement of the nucleosomal DNA and hence of the attachment site.

Conformational dynamics within Chd1-nucleosome complexes.

To determine the positions of tandem chromodomains, ATPase domain and DNA-binding domain of Chd1 within the Chd1-nucleosome complex, I measured with the help of Monika

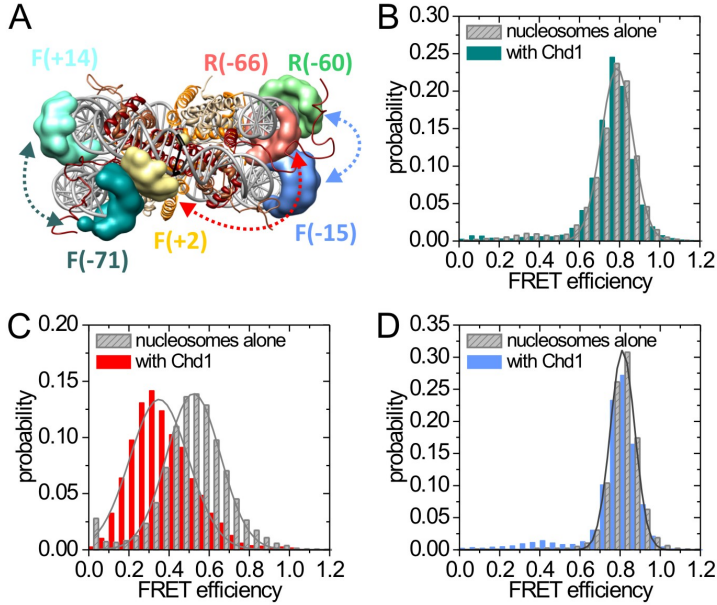


Figure 4.4.2.: Validation of the structure of nucleosomal DNA within Chd1-nucleosome complexes. (A) Schematic illustration of the smFRET measurements designed to test whether the structure of the nucleosome particle is a good model for the Chd1-nucleosome complex with respect to the position of nucleosomal DNA. (B-D) smFRET experiments were performed from F(-71)-Alexa647 to F(+14)-Tamra (B, dark cyan), from F(+2)-Tamra to R(-66)-Alexa647 (C, red) and from F(-15)-Tamra to R(-60)-Alexa647 (D, light blue) in the absence (gray histograms) and in the presence (colored histograms) of bound Chd1. The latter two measurements were performed by Monika Holzner.

Holzner smFRET for all combinations of satellite and antenna pairs yielding information for a total of 39 distances (Figure 4.4.3 and supplemental data A.3). One exemplary set of smFRET measurements from all satellites to a position on the DNA-binding domain, namely the data for NPS localization of antenna Chd1-C1149, is shown in Figure 4.4.4.

Notably, the FRET efficiency distributions of all measurements between satellites and antenna Chd1-C1149 were best described by two Gaussians suggesting two alternative locations of antenna C1149 relative to the nucleosome core particle. The Gaussian fits revealed major FRET populations comprising 65-81 % of the data and subpopulations at higher (F(-64), F(-15), F(+2) and F(+14)) or lower (R(-60) and R(-66)) FRET efficiencies comprising 19-35 % of the data. A closer examination of smFRET time trajectories revealed dynamic switching between the two FRET populations for about 2-10 % of all complexes. An exemplary dynamic trace is shown for the measurement between satellite F(-15)-Tamra and Chd1-C1149-Alexa647 together with the corresponding smFRET distribution in Figure 4.4.5A. Dynamic switching between a low FRET state at approximately

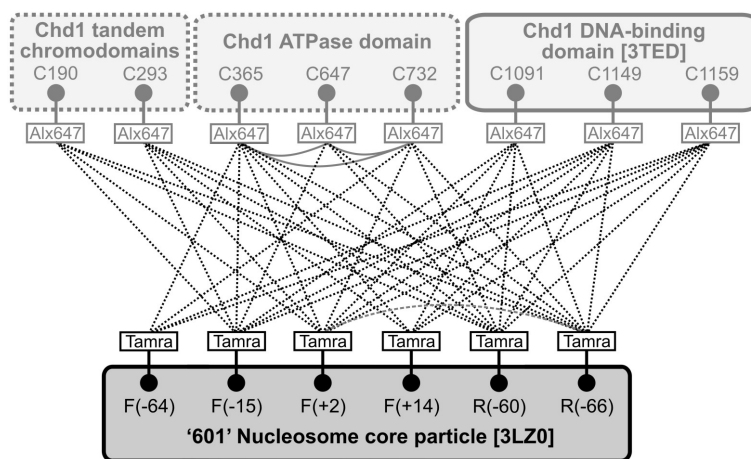


Figure 4.4.3.: Schematic representation of FRET network for NPS docking analysis of the Chd1-nucleosome complex. All satellites and antennas are shown in their respective reference frame together with the attached dye molecules. FRET efficiencies were measured between pairs of satellite and antenna dyes (dotted black lines, $n = 39$) and in between satellites (dashed gray line, $n = 1$), respectively. The location and orientation of tandem chromodomain antennas C190 and C293, ATPase antennas C365, C647 and C732 as well as of the reference frame of the DNA-binding domain with the three attached antennas C1091, C1149 and C1159 (gray-hued) relative to the nucleosome core coordinate system (dark-hued) were to be determined by global NPS analysis [22]. Distance constraints in between antennas attached to the ATPase domain as inferred from the available structural data are indicated as solid gray lines.

20 % and a high FRET state at 60 % on the order of seconds is visible. The observed dynamics demonstrated that the two different FRET populations in FRET efficiency histograms did not arise from two different population of Chd1-nucleosome complexes, but from single complexes interconverting between two different conformations. Next, I explored whether other parts of the DNA-binding domain, namely antennas Chd1-C1091 and Chd1-C1159, also showed dynamics relative to the nucleosome core particle. Indeed, smFRET efficiency distributions for all measurements from the satellites to antennas Chd1-C1091 and Chd1-C1159 were bimodal with one major population comprising 66-81 % of the data and a minor supopulation visible as a shoulder and comprising 19-34 % of the data (Figure A.3.1 and A.3.2) and again, dynamic interconversion between the two FRET populations was observed. One dynamic smFRET time trace is shown for the measurement between satellite F(+14)-Tamra and antenna Chd1-C1159-Alexa647 in Figure 4.4.5B, revealing dynamic transitioning between two states at about 20 and 50 % FRET on the time scale of seconds. The dwell time of complexes in the high FRET state at 50 % FRET is shorter consistent with the lower probability (20 %, see Table A.3.1) of this FRET state in the framewise FRET efficiency distribution. Conformational dynamics of the DNA-binding domain of Chd1 relative to the nucleosome core particle are consistent with the earlier observation that the long extranucleosomal DNA on the nu-

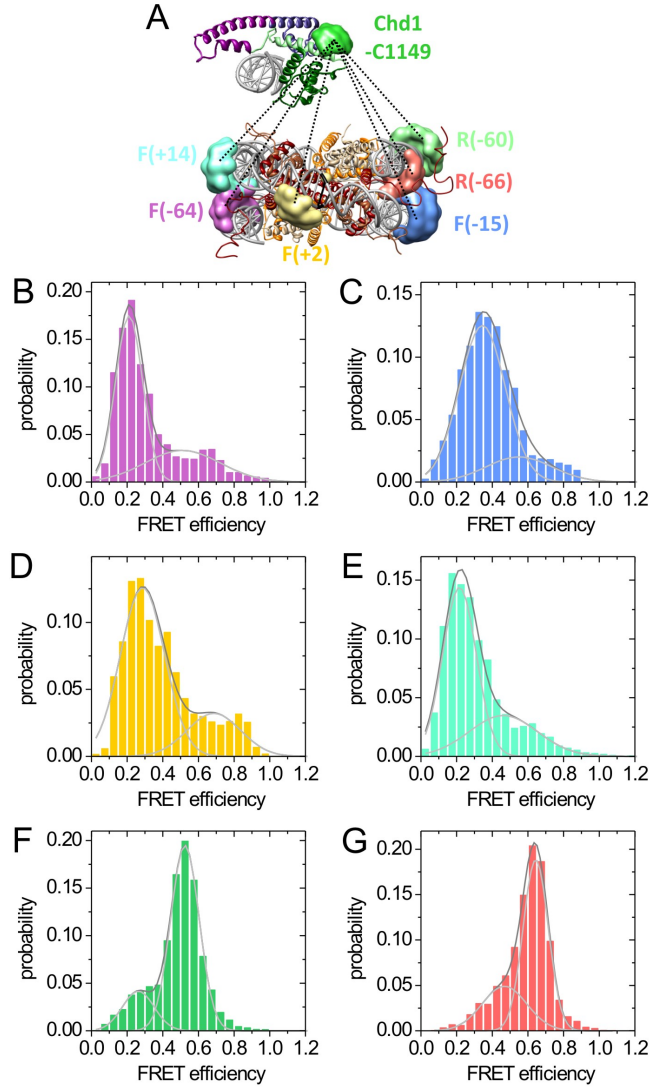


Figure 4.4.4.: Positioning of Chd1 relative to the nucleosome core particle. (A) Schematic illustration of an exemplary set of smFRET measurements from all satellites to a position on the DNA-binding domain, namely the data for NPS localization of antenna Chd1-C1149. The images of nucleosome core particle and Chd1 DNA-binding domain containing the dye position priors were prepared in Chimera [151]. The measurements are indicated by dotted black lines and the respective FRET efficiency histograms are shown in the color of the corresponding satellite (B-G) (magenta for satellite F(-64), light blue for F(-15), yellow for F(+2), aquamarine for F(+14), light green for R(-60), light red for R(-66)). All FRET efficiency histograms were best described with two Gaussians since two populations were visible, one major population comprising 65-81 % of the data and a subpopulation comprising 19-35 % of the data (light gray: individual fits, dark gray: combined fits).

4 Single molecule FRET studies of nucleosome remodeling by Chd1

cleosome entry site itself undergoes dynamic conformational changes in the presence of Chd1 and AMP-PNP (section 4.3.1, Figure 4.3.6). This suggests, that Chd1 binds via its DNA-binding domain to the long linker DNA and induces dynamics thereby moving together with the DNA.

Conformational dynamics were not limited to the DNA-binding domain. Also in most smFRET measurements between satellites and antennas attached to the chromodomains or the ATPase domain of Chd1, bimodal (or even trimodal in the measurement from F(-15)-Tamra to Chd1-C293-Alexa647) FRET efficiency distributions were obtained (Supplemental data, Figures A.3.3, A.3.4, A.3.5, A.3.6, A.3.7) and dynamic switching between the different FRET populations was observed in a small fraction of complexes. Exemplary dynamic smFRET time trajectories are shown for the measurement from chromodomain antenna Chd1-C293 to satellite R(-66) and for the measurement from ATPase antenna Chd1-C365 to satellite F(+14) (Figure 4.4.5C and D, respectively). These results indicate that not only the DNA-binding domain of Chd1, but also tandem chromodomains and ATPase domain move relative to the nucleosome core particle even in the presence of AMP-PNP and therefore capture different conformations.⁴

NPS results for Chd1 DNA-binding domain, tandem chromodomains and ATPase domain

Global NPS analysis emanates from a static model and hence, each fluorophore is assumed to capture a static position in the complex of interest [22]. Therefore, the intrinsically dynamic Chd1-nucleosome complexes can only be analyzed by treating each intermediate conformational state as static. In this work I will focus on the architecture of the Chd1-nucleosome complexes in the major conformation corresponding to the main population in smFRET efficiency histograms. Analysis of Chd1-nucleosome complexes in the alternative conformation corresponding to the side peaks in the FRET efficiency histograms as well as of the structural dynamics in between both conformations will be the subject of future research.

For global NPS analysis of the Chd1-nucleosome complex, in addition to acquiring smFRET data for all combinations of satellites and antennas, steady-state fluorescence anisotropies of the donor and acceptor dyes for all attachment sites were measured and the isotropic Förster distance for each acceptor and donor dye pair was experimentally determined (Table A.3.1 and section 4.6.17). The structure of the DNA-binding domain of Chd1 within a Chd1-nucleosome complex was assumed to be unchanged compared to its x-ray crystallographic structure in complex with DNA [150]. Therefore, NPS docking [22] was applied and the rigid structure of the Chd1 DNA-binding domain in complex

⁴I should mention that the majority of dynamic FRET trajectories observed for all Chd1 domains consisted of a single transition and multiple back and forth switching as shown in Figure 4.4.5 was observed only in a minority of traces. The observation of multiple transitions requires a long enough duration of FRET trajectories, which was mostly not given due to fast photobleaching of the acceptor dye Alexa647.

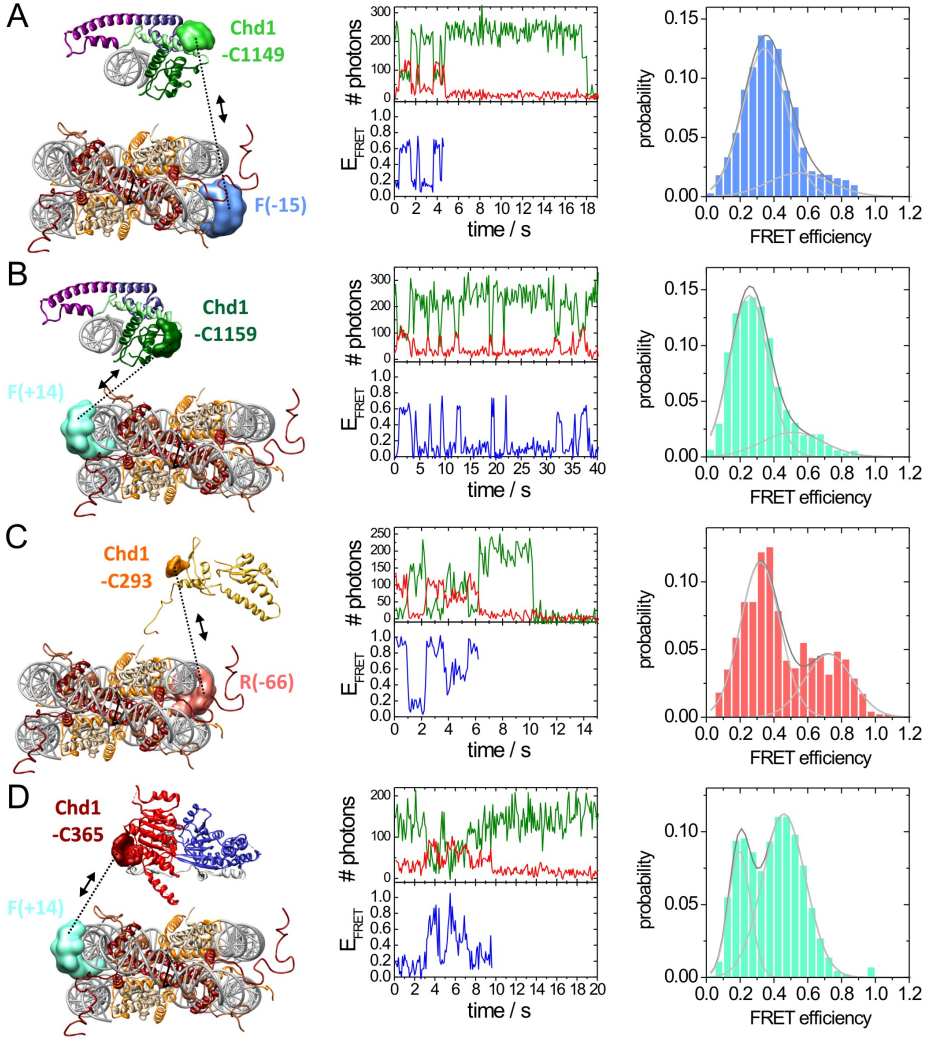


Figure 4.4.5.: Conformational dynamics within Chd1-nucleosome complexes. Exemplary dynamic smFRET time traces are shown for antennas on all three Chd1 domains (center) together with the corresponding framewise FRET efficiency histogram (right) and with a cartoon illustrating the measured distance (left). (A) smFRET dynamics in the measurement from DNA-binding domain antenna Chd1-C1149-Alexa647 to satellite F(-15)-Tamra. (B) smFRET dynamics in the measurement from DNA-binding domain antenna Chd1-C1159-Alexa647 to satellite F(+14)-Tamra. (C) smFRET dynamics in the measurement from chromodomain antenna Chd1-C293-Alexa647 to satellite R(-66)-Tamra. (D) smFRET dynamics in the measurement from ATPase antenna Chd1-C365-Alexa647 to satellite F(+14)-Tamra.

4 Single molecule FRET studies of nucleosome remodeling by Chd1

with DNA (pdb ID: 3TED) was defined as reference frame that was to be localized and oriented relative to the nucleosome core particle during global docking NPS analysis (see Experimental Procedures 4.6.17). Antennas attached to the DNA-binding domain, namely Chd1-C1091, -C1149 and -C1159, were assigned to this reference frame. The chromodomain antennas, Chd1-C190 and -C293, were treated as independent dye molecules without assignment to a reference frame, since the available structures of Chd1 tandem chromodomains [147, 142] are likely to change when being part of a Chd1-nucleosome complex (see section 1.2.5). Similarly, antennas attached to the Chd1 ATPase domain, Chd1-C365, -C647 and -C732, were treated as independent dyes without reference frame, however distance constraints in between the antenna attachment points were applied since the magnitude of conformational change within the ATPase domain during binding to a nucleosome could be well estimated based on available structural information [147, 148, 159, 260] (Experimental Procedures, section 4.6.17). In the global NPS docking analysis, the complete data set consisting of mean FRET efficiencies, steady-state fluorescence anisotropies, and isotropic Förster distances was used to simultaneously infer positions and orientations of all antennas and the corresponding reference frames within the coordinate system of the nucleosome core particle. As a result of Bayesian parameter estimation, the marginal three-dimensional position probability density function of each antenna and of any desired position within the docked reference frames were obtained.

Figure 4.4.6A shows the NPS results for the Chd1 DNA-binding domain (DBD) regarding the position densities (68 % credibility) of antenna attachment points C1091, C1149 and C1159, as extrapolated from the position density of the Chd1-DBD reference frame, relative to the coordinate system of the nucleosome core particle [99, 256]. All DBD densities are located above the nucleosome close to the nucleosome entry site. Positions C1149 in the helical linker 2 (HL2) region (light green) and C1159 in the SLIDE domain are situated above the first basepairs of the nucleosomal DNA at the entry site and adjacent histone H3, whereas position C1091 at the end of helical linker 1 (HL1) is shifted further away from the nucleosome in the direction indicated by the entering DNA and is localized close to the N-terminal tail of histone H3. The densities were used to build a preliminary model of the Chd1-DBD bound to the nucleosome as part of the complete Chd1-nucleosome complex. Therefore, the DBD-DNA structure (Figure 4.4.6A, gray box) was moved as a rigid body such that the antenna attachment points were situated in their respective densities, and such that the DNA stretch bound by the DBD was able to connect with the nucleosomal DNA at the entry site. Moreover, the earlier results were also taken into account that nucleosomal DNA around position R(-66) may be slightly displaced in the Chd1-nucleosome complex compared to its position in the nucleosome crystal structure, whereas nucleosomal DNA at position R(-60) is unchanged (see Figure 4.4.2). In the resulting model (Figure 4.4.6B), the first 9 bp of nucleosomal DNA at the entry site are bent upwards such that the DNA enters the nucleosome at an angle of about 30 ° relative

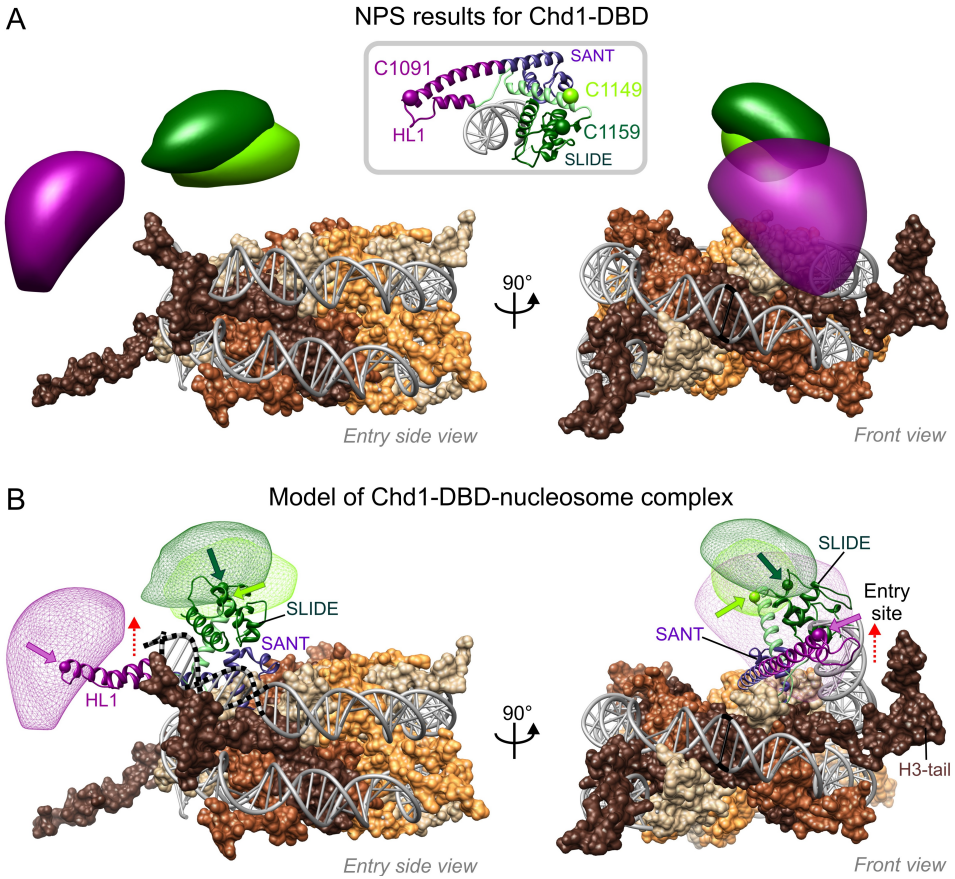


Figure 4.4.6.: NPS results and preliminary model of Chd1 DNA-binding domain (DBD) bound to a nucleosome as part of a Chd1-nucleosome complex. (A) Global NPS results of Chd1-DBD antenna attachment points C1091 (magenta), C1149 (light green) and C1159 (dark green) are shown as 68 % credibility volumes in the coordinate system of the nucleosome (pdb IDs: 3LZ0 (601 DNA) and 1KX5 (histones with tails), [99, 256]) in the entry side view (left) and the front view (right). The DNA is presented as gray ribbon (dyad, black) and the histones are depicted as surfaces (H3, dark brown; H4, sienna; H2A, beige; H2B, light brown). In the gray box, the DBD-DNA structure (pdb ID: 3TED) is shown with the antenna attachment points presented as spheres in the respective color. Helical linker 1 (HL1), magenta; SANT, purple; SLIDE, dark green; helical linker 2 (spacer helix), light green. (B) Preliminary model of the Chd1-DBD bound to a nucleosome based on global NPS docking results. The antenna attachment points of C1091, C1149 and C1159 (same color coding as in panel A) are shown as spheres (indicated by arrows) together with the corresponding NPS densities as meshed credible volumes at 68 % probability. (left, entry side view; right, front view). Modeling required the first 9 bp of the nucleosomal DNA at the entry site to be bent upwards as indicated by black, dotted lines in the entry side view and red, dotted arrows in both views.

4 Single molecule FRET studies of nucleosome remodeling by Chd1

to the dyad axis. The DBD binds the extranucleosomal DNA right at the nucleosome entry site slightly above the nucleosome, with the HL1 motive pointing away from the nucleosome, the SLIDE domain facing upwards and the SANT domain facing down to the histone octamer surface, particularly to histone H3 at the entry site.

The NPS results for the antennas on the Chd1 tandem chromodomains in the coordinate system of the nucleosome core particle are presented in Figure 4.4.7A. The density of antenna C190 in chromodomain 1 has an elongated shape and is localized underneath the nucleosome core at the entry site next to the protruding N-terminal tail of histone H4. The density of chromodomain 2 antenna C293 is similarly elongated and is situated in front of the nucleosome between entry site and dyad. Both densities were used to build a preliminary model of the Chd1-tandem chromodomains relative to the nucleosome as part of the complete Chd1-nucleosome complex. Even though the chromodomain antennas were treated as individual dyes without assignment to a reference frame during the global NPS analysis, for the purpose of modeling the conformation of tandem chromodomains was assumed to be rigid and unchanged compared to its conformation in the Chd1 chromodomain-ATPase crystal structure [147]. The tandem chromodomains were then moved such that the accessible volumes of the two chromodomain antennas were situated in their respective NPS densities (Figure 4.4.7B). No information existed about the position of the chromo-wedge motive connecting both chromodomains and therefore the orientation of the chromodomains was completely unknown since the geometry of chromodomain antenna densities allowed any orientation of the chromo-wedge. However, previous mutational and structural studies suggested that yeast Chd1 chromodomains interact with the H4 tail [142], in contrast to human Chd1 chromodomains that was shown to target trimethylated lysine 4 of the tail of histone H3 [146]. Therefore, I rotated the chromo-wedge motive such that it locates near the bottom of the nucleosome close to the protruding H4 tail. The overall proximity of the modeled tandem chromodomains to the nucleosome entry region, from where the tails of histones H3 and H4 protrude, is consistent with previous structural and mutational studies [146, 142, 143, 261] and proposes that the chromodomains of human Chd1 might capture a similar position thereby interacting with trimethylated H3K4.

Finally, the global NPS densities for the Chd1 ATPase antennas all localized above the nucleosome as visible in Figure 4.4.8A. The density of antenna C365 in ATPase lobe 1 is positioned right on top of the center of the histone octamer surface whereas the density for antenna C732 in ATPase lobe 2 is positioned above the front of the nucleosome around the dyad. In contrast to these two relatively well defined densities, antenna C647 in ATPase lobe 2 had a large NPS density above the whole front half of the nucleosome. The large uncertainty of this density is the result of using only two distance measurements for positioning (Supplemental Figure A.3.6), since most measurements for antenna C647

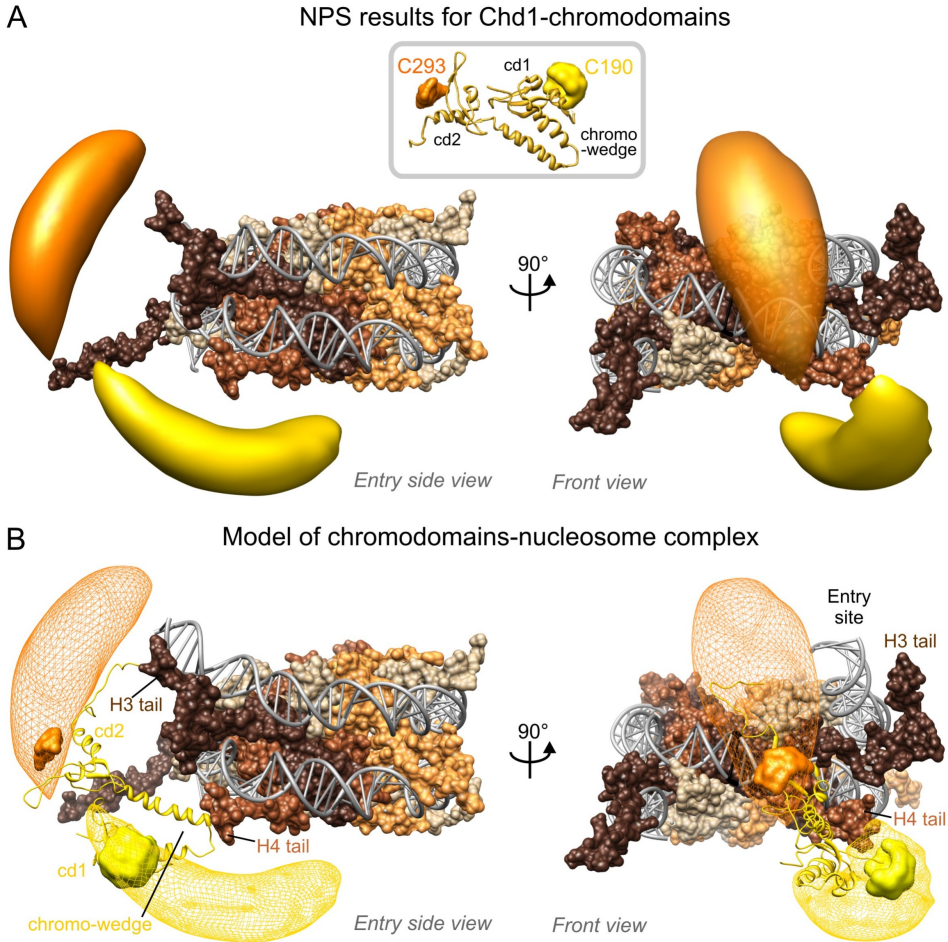


Figure 4.4.7.: NPS results and preliminary model of Chd1 tandem chromodomains bound to a nucleosome as part of a Chd1-nucleosome complex. (A) Global NPS results of Chd1-chromodomain antennas C190 (yellow) and C293 (orange) are shown as 68 % credibility volumes in the coordinate system of the nucleosome (pdb ID: 3LZ0 (601 DNA) and 1KX5 (histones with tails)) in the entry side view (left) and the front view (right). The DNA is presented as gray ribbon (dyad, black) and the histones are depicted as surfaces (H3, dark brown; H4, sienna; H2A, beige; H2B, light brown). The first 9 bp of the nucleosomal DNA at the entry site are bent upwards as a result of modeling of the DNA-binding domain (see Figure 4.4.6). In the gray box, the Chd1 double-chromodomain structure (pdb ID: 3MWY) is shown together with the accessible volumes of the antenna dyes as solid densities in the respective color (cd1, chromodomain 1; cd2, chromodomain 2). (B) Preliminary model of the Chd1-tandem chromodomains bound to a nucleosome based on global NPS results. The accessible volumes of chromodomain antennas C190 and C293 (same color coding as in panel A) are shown as solid densities together with the corresponding NPS densities as meshed credible volumes at 68 % probability. (left, entry side view; right, front view). Flexible N-terminal tails of histones H3 and H4 protruding from the nucleosome core are in close proximity to the modeled chromodomains and the chromo-wedge (see gray box panel A).

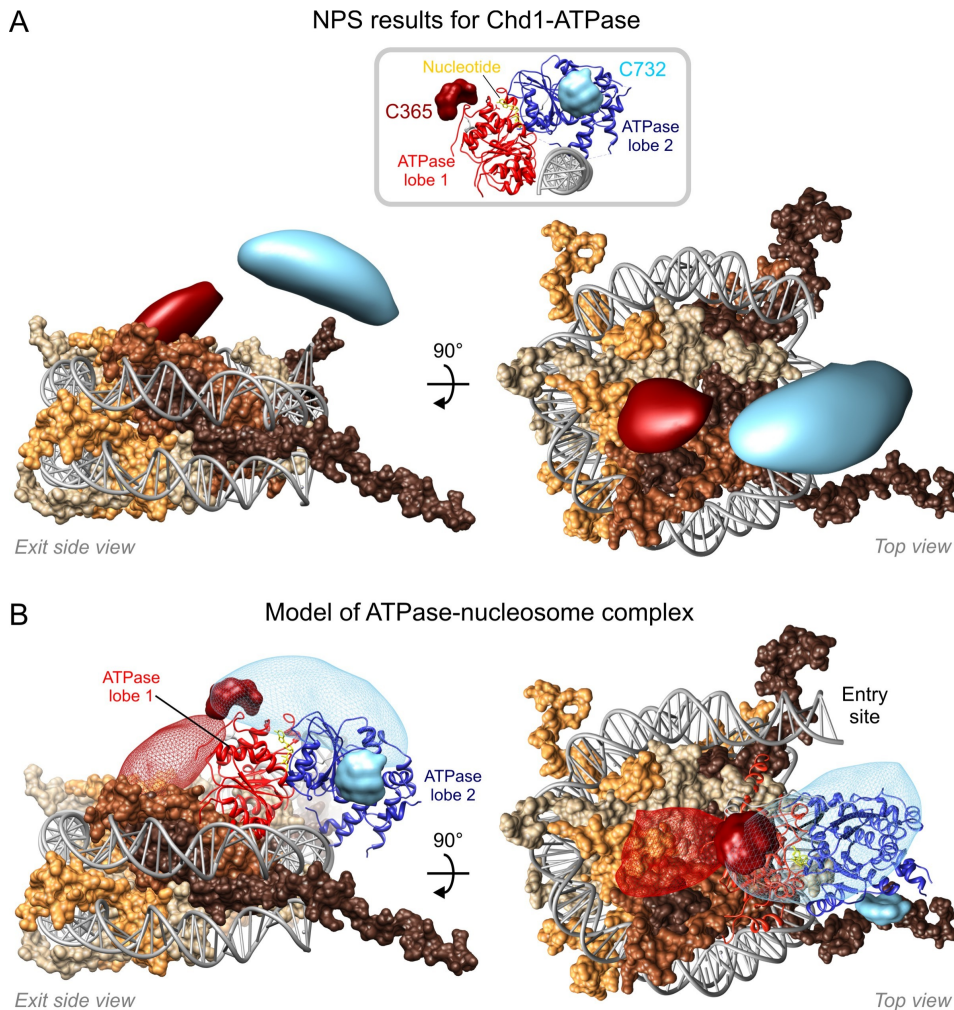


Figure 4.4.8.: NPS results and preliminary model of Chd1 ATPase domain bound to a nucleosome. (A) Global NPS results of Chd1-ATPase antennas C365 (dark red) and C732 (light blue) are shown as 68 % credibility volumes in the coordinate system of the nucleosome (pdb ID: 3LZ0 (601 DNA) and 1KX5 (histones with tails)) in the exit side view (left) and the top view (right). The DNA is presented as gray ribbon (dyad, black) and the histones are depicted as surfaces (H3, dark brown; H4, sienna; H2A, beige; H2B, light brown). In the gray box, the modeled closed Chd1 ATPase structure in complex with DNA and nucleotide (pdb ID: 3MWY, [147] modified as outlined in Figure A.3.8) is shown together with the accessible volumes of the antenna dyes as solid densities in the respective color. (B) Preliminary model of the Chd1-ATPase domain bound to a nucleosome based on global NPS results of antennas C365 and C732. The accessible volumes of ATPase antennas C365 and C732 (same color coding as in panel A) are shown as solid densities together with the corresponding NPS densities as meshed credible volumes at 80 % probability. (left, exit side view; right, top view).

yielded extremely low statistics. Therefore, the density of antenna C647 was disregarded during the following modeling and is not shown in Figure 4.4.8.

Prior to modeling the Chd1 ATPase domain relative to the nucleosome based on the NPS densities, the following considerations were made: In the existing crystal structure, the Chd1 ATPase domain captures an open, inactive conformation, in which both ATPase lobes are splayed relatively far apart and the ATPase DNA-binding surface is occluded by the chromo-wedge motive [147] (Figure 1.2.4). However, I assume that in the Chd1-nucleosome complex, the ATPase domain captures a closed, catalytically active conformation as it has been observed for the zebrafish Rad54 ATPase structure with bound nucleotide [260] or for the drosophila RNA helicase Vasa bound to AMP-PNP and to RNA template [262]. This assumption is supported by the fact that the Chd1 ATPase densities are positioned on top of the nucleosome, whereas the densities of the Chd1 chromodomains are located in front of and below the nucleosome, indicating the disruption of the interaction between ATPase and chromo-wedge. Moreover, AMP-PNP was present in all smFRET positioning experiments, which is thought to facilitate the closure of the ATPase lobes. Therefore, a model of the closed, catalytically competent Chd1 ATPase domain in complex with DNA was constructed by structural comparison with the closed zebrafish Rad54 ATPase structure [260] and the *Sulfolobus solfataricus* Rad54 ATPase structure in complex with DNA [159] (Supplemental Figure A.3.8). The closed ATPase domain bound to DNA, shown in the gray box in Figure 4.4.8, was then moved relative to the nucleosome such that the accessible volumes of the antennas C365 and C732 overlapped with the respective NPS densities and such that the DNA, bound by the ATPase matched the nucleosomal DNA. The resulting model of the Chd1 ATPase domain bound to the nucleosome core particle is shown in Figure A.3.8B. The ATPase domain is situated on top of the nucleosomal DNA between the dyad (SHL0) and SHL+1, thereby binding the minor groove in between. ATPase lobe 1 faces the center of the nucleosome, whereas ATPase lobe 2 faces outwards. Notably, the accessible volumes of both antennas only minimally overlap with the respective NPS densities, even though the densities are presented at 80 % instead of 68 % credibility. However, forcing a stronger overlap between NPS densities and accessible volumes either led to a major clash of ATPase lobe 1 with the histone octamer surface or to a displacement of the nucleosomal DNA around the dyad, which is expected to be very unlikely. Due to this incongruity, I tested whether the assumption that the ATPase motor captures a closed conformation in the Chd1-nucleosome complex was wrong. I tried to model the ATPase domain in diverse open conformations [147, 159] relative to the nucleosome into the NPS densities, however no better result could be achieved. One explanation for the unsatisfactory overlap between ATPase antenna C365 and the corresponding density is that the attachment point C365 might be displaced during binding of Chd1 to the nucleosome compared to its position relative to ATPase lobe 1 in the ATPase-chromodomains complex structure. Since C365 is located within a flexible loop connecting ATPase lobe 1 with the tandem chromodomains,

4 Single molecule FRET studies of nucleosome remodeling by Chd1

a rearrangement is likely, especially since the chromodomains are assumed to detach and move away from the ATPase domain. In turn, position C732 is located within a helix motive and is not expected to alter markedly during Chd1 binding to the nucleosome. In summary, more information is required regarding the position of the ATPase domain relative to the nucleosome and regarding the conformation of both ATPase lobes relative to each other in order to solidify the preliminary model presented here.

4.4.2. Discussion

Here, we determined the position of the three constituting domains of chromatin remodeler Chd1 relative to the nucleosome core particle in a Chd1-nucleosome complex by using smFRET experiments, global NPS analysis as well as x-ray crystallographic data. The results of global NPS analysis together with x-ray crystallographic information about the nucleosome core particle as well as the domains of Chd1 allowed the construction of a preliminary model of the overall architecture of a Chd1-nucleosome complex (Figure 4.4.9). In this model, the DNA-binding domain binds extranucleosomal DNA just outside the nucleosome entry site slightly above the nucleosome and the first turn of nucleosomal DNA at the entry site is bent upwards by about 30° . The tandem chromodomains are situated below the DNA-binding domain in front of and below the nucleosome near the tails of histones H3 and H4, protruding from the nucleosome core. Finally the ATPase domain is bound from the top to the nucleosomal DNA between dyad and SHL+1 with ATPase lobe 1 being located above the histone octamer surface and ATPase lobe 2 facing away from the nucleosome. The Chd1-nucleosome complex showed intrinsic dynamics, which directly explains the inability of x-ray crystallographers to capture this complex in a crystal structure.

In the following, I will compare the architecture of a Chd1-nucleosome complex presented here with previous biochemical and structural results for Chd1 and other remodeling machines, especially members of the ISWI family. The association of the DNA-binding domain of Chd1 with a short stretch of the long extranucleosomal DNA close to the nucleosome entry site is consistent with earlier observations that the DNA-binding domain requires the presence of extranucleosomal DNA for a stable association with the nucleosome and that this association with the linker DNA determines the direction of nucleosome sliding towards the center of the DNA [152, 149]. However, the observed orientation and position of the SANT-SLIDE motive relative to the nucleosome core particle and nucleosomal DNA is distinct from the orientation and position reported for the SANT-SLIDE motives of the HAND-SANT-SLIDE domain of ISWI remodelers [167, 154]. Hence, the SANT-SLIDE motive shared between the ISWI and Chd1 family of remodelers might function in a distinct way in each remodeler family. The Chd1 DNA-binding domain appears to occupy a shorter stretch of extranucleosomal DNA closer to the nucleosome than it was shown for the DNA-binding domain of IWSI [167, 154], which might explain the previous observation that Chd1 assembles chromatin with a nucleosome repeat length of 162 bp, which is shorter than the nucleosome repeat length of 175 bp introduced by the ISWI-type remodeler ACF [127].

Moreover, the overall proximity of the modeled tandem chromodomains to the nucleosome entry region, from where the tails of histones H3 and H4 protrude, is consistent with previous structural and biochemical studies [146, 142, 143, 261], which showed that the tandem chromodomains target these histone tails. It suggests that the chromodomains of human

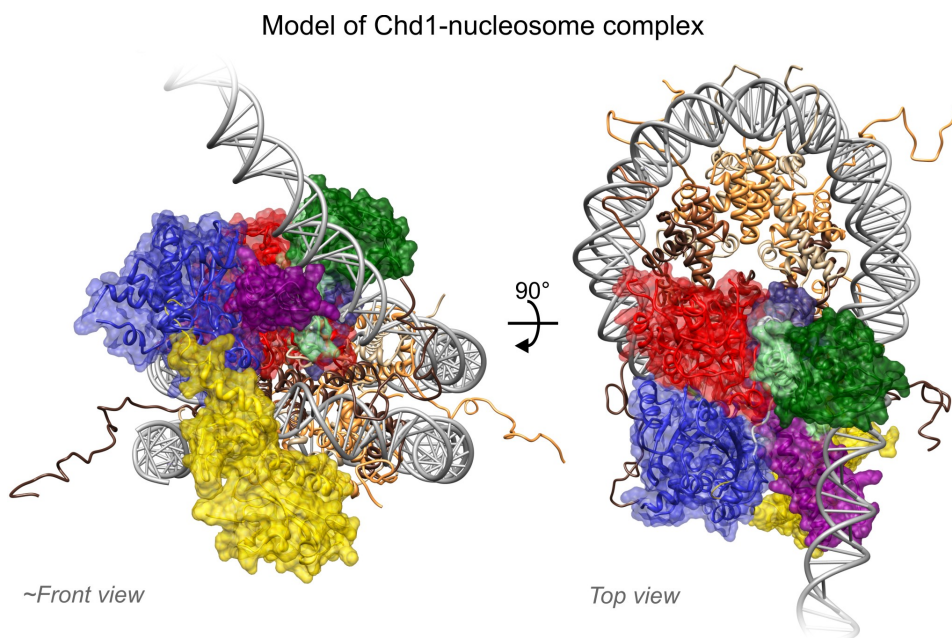


Figure 4.4.9.: Preliminary model of the overall architecture of a Chd1-nucleosome complex.

The nucleosome is presented in the front view slightly rotated clockwise to better show the entering DNA (left) and in the top view (right). DNA and histones are presented as ribbon (DNA, dark gray; H3, dark brown; H4, sienna; H2A, beige; H2B, light brown), whereas the different Chd1 domains are depicted as semi-transparent surfaces enclosing the ribbon representation. The same color coding for the individual structural motives of Chd1 is used as in previous figures with the chromodomains in yellow, ATPase lobe 1 in red, ATPase lobe 2 in blue and the DNA binding domain in magenta (HL2), purple (SANT), light green (HL1, spacer helix) and dark green (SLIDE). The long extranucleosomal DNA at the nucleosome entry site was modeled by straightly extending the DNA bound by the DNA-binding domain.

Chd1 might capture a similar position, since they were shown to cooperate to recognize methylated lysine 4 of histone H3 [146]. Responsible for this interaction are two aromatic residues in the interface between both chromodomains, one of which, tryptophane 67, is missing in *S.cerevisiae* Chd1, leading to a lack of the interaction between H3K4me3 and Chd1 chromodomains in yeast. In turn, the chromodomains of yeast Chd1 have been proposed to associate with the tail of histone H4 [147], but it remains unclear whether a specific histone mark within the H4 tail is recognized. Since only two positions within the chromodomains were localized by NPS analysis in this work, the orientation of chromodomains and the position of the chromo-wedge motive were unknown and had to be guessed during modeling based on previous information. Therefore, the position of the chromo-wedge motive should be additionally determined by NPS to solidify the model of the chromodomains relative to the nucleosome.

The location of the Chd1 ATPase domain between the dyad and SHL+1 presented here is close to the position two turns away from the dyad (SHL+2), where the ATPase domain of

ISWI remodelers interact with the nucleosomal DNA as shown by biochemical crosslinking and electron microscopy [164, 165, 166, 167, 172]. However, there is a notable shift of about 10-15 base pairs between both positions and it remains to be shown, whether this deviation is real and indicates differences in the detailed remodeling mechanism of the two remodeler families or whether it is an artifact resulting from wrong assumptions during modeling or from an insufficient accuracy of the NPS results. The fact that introduction of single-stranded gaps at SHL2 blocks sliding of the gapped sites towards the dyad not only in the case of ISWI and SWI/SNF remodelers [166, 165, 168, 169], but also in the case of Chd1 [152] suggests that the ATPase domains of both remodeler families bind at a similar location. Hence, further experiments should be performed to solidify the position of the Chd1 ATPase domain relative to the nucleosome. Moreover, experiments should investigate the conformation of the ATPase lobes relative to each other in the Chd1-nucleosome complex, since a change in conformation compared to the closed ATPase conformation used here for modeling could account for discrepancies between the NPS model and previous biochemical results.

The Chd1-nucleosome model presented here, even though preliminary, has implications for the mechanism of remodeling. The DNA-binding domain is located in close proximity to the ATPase domain and can directly contact both ATPase lobes via its long helix containing HL1 and SANT domain. This interaction between both domains could enable a direct coupling of ATP hydrolysis by the ATPase to the breakage of histone-DNA contacts at the nucleosome entry site by the DNA-binding domain. Moreover, the relative position of tandem chromodomains and ATPase lobes directly demonstrates how the chromo-wedge motive, which inhibits the ATPase motor in a ATPase-chromodomains complex, opens up when it encounters a nucleosome to allow for binding of the ATPase to the nucleosomal DNA, as proposed previously [147]. The chromodomains were suggested to help the ATPase motor to discriminate between a nucleosome and a naked DNA substrate. The scenario can be imagined in which the double-chromodomains of Chd1 recognize and bind the H4 tail of a nucleosome (or trimethylated lysine 4 in H3 tail in case of human Chd1), thereby anchoring the remodeler to the nucleosome substrate and simultaneously freeing the ATPase motor for subsequent interaction with nucleosomal DNA. The role of the DNA-binding domain would then be to support the association of Chd1 with the nucleosome through interaction with extranucleosomal DNA and to determine the binding orientation by preferentially binding to the longer available linker DNA of a nucleosome.

Finally, the model presented here is only preliminary and experiments to finalize and solidify the positions of the different Chd1 domains are still ongoing. One general problem of positioning a remodeler relative to a nucleosome by NPS analysis is the limitation of satellite positions to the nucleosome core particle with its disk like geometry. As a result,

the direction perpendicular to the nucleosome dyad axis cannot well be explored. Moreover, if part of the remodeler such as the DNA-binding domain binds further out of the nucleosome core particle to the extranucleosomal DNA, it cannot be positioned by using satellites on the nucleosome core particle. To increase the accuracy of the NPS densities of Chd1 antennas, it will be helpful to add information about the relative positions of Chd1 domains by performing smFRET measurements in between different antennas. Therefore, double-labeled Chd1 mutants are required. To better localize parts of the remodeler situated in close proximity to the extranucleosomal DNA, positions on the extranucleosomal DNA can be used as additional antennas, incorporated into the FRET network through smFRET measurements to antennas on the individual Chd1 domains and to satellites on the nucleosome core particle. In this way, the extranucleosomal DNA itself could be positioned relative to Chd1 and the nucleosome. Moreover, combination of the NPS approach with structural data from small angle x-ray scattering, which can provide information about the overall shape of the Chd1-nucleosome complex, could further strengthen a model of the architecture of the Chd1-nucleosome complex.

4.5. Summary and Outlook

In the study presented in this chapter, I used single molecule FRET to provide new insights into two different aspects of nucleosome remodeling by Chd1. First, I explored the structural dynamics of individual nucleosomes during binding and remodeling by individual Chd1 remodelers in real-time. Time resolved smFRET data directly revealed gradual and bidirectional translocation of nucleosomal DNA by continuously bound Chd1 on a time scale of hundreds of milliseconds to seconds and allowed me to propose a remodeling mechanism applied by Chd1, which involves formation and propagation of a DNA loop. This study presents the first real-time observation of remodeling by individual Chd1 remodeling machines.

Second, I used smFRET combined with NPS analysis and x-ray crystallographic data to determine the three-dimensional architecture of a Chd1-nucleosome complex prior to remodeling. The Chd1-nucleosome complexes, even though trapped in a non-active conformation by the presence of AMP-PNP, showed intrinsic dynamics. Both, the different remodeler domains as well as the nucleosomal DNA at the entry site were moving relative to the nucleosome core particle. These dynamics directly explain why traditional structural methods such as x-ray crystallography have been hindered in structurally investigating Chd1-nucleosome complexes and demonstrate the power of the single molecule methods used here.

The study presents a significant progress towards understanding the mechanism of nucleosome remodeling by Chd1, however many open questions remain. The proposed Chd1 remodeling mechanism as well as the proposed architecture of the Chd1-nucleosome complex are both preliminary models and future experiments should aim for solidifying the models and diminishing existing uncertainties.

Concerning the Chd1-nucleosome architecture, the accuracy of the NPS results should be improved by further connecting the present FRET network through smFRET measurements in between Chd1 antenna positions using dual-labeled Chd1 bound to nucleosomes. Further, information about the relative orientation of the individual FRET dye pairs should be included by measuring FRET anisotropies, as it was shown to substantially increase the accuracy of NPS [22]. Moreover, it will be helpful to perform three color smFRET experiments [263], in which the nucleosome is labeled at two deliberate positions on the nucleosomal DNA or on histones, but also the remodeler Chd1 contains one fluorophore. In this way, distances and dynamics within the nucleosome can be observed simultaneously to distances and dynamics of the remodeler relative to the nucleosome. Therefore the presence and position of one or more remodelers as well as the conformational state of the nucleosome can be investigated simultaneously. Another profitable endeavor would be the combination of NPS analysis with other low resolution methods for structure determination such as small angle x-ray scattering (SAXS), pulsed dipolar electron spin resonance (ESR) or biochemical crosslinking experiments. Whereas SAXS can provide information about the overall shape of a macromolecule, ESR delivers long

4 Single molecule FRET studies of nucleosome remodeling by Chd1

range distance information in solution and biochemical crosslinking provides information about residues in close proximity. All three approaches are already applied in the laboratory of our collaborator Tom Owen-Hughes, though only to Chd1 alone in the absence of nucleosomes. It will be exciting to see how the different methods can be combined in a hybrid approach to yield a more detailed model of the architecture of a Chd1-nucleosome complex.

Concerning the direct observation of Chd1 remodeling dynamics, it was shown that a measurement time resolution of 200 ms is not sufficiently high to observe all transitions. Therefore, smFRET experiments of Chd1 remodeling dynamics have to be repeated at a higher time resolution. Moreover, it remains to be shown if Hidden Markov Modeling is the appropriate method to analyze the dynamic smFRET traces of nucleosome remodeling.

Furthermore, it will be of great benefit to apply three color smFRET experiments in order to connect conformational changes of the remodeler with consequential structural rearrangements within the nucleosome. Along this line, it would be very helpful to use the information gained from NPS analysis about the architecture of stalled Chd1-nucleosome complexes in order to put forward hypotheses for possible conformational changes within the complex during remodeling. Deliberate label positions could then be designed and the hypotheses could be tested by means of time-resolved smFRET experiments.

All experiments presented here were performed with nucleosomes containing the strongest known nucleosome positioning sequence 601 [111]. This sequence is characterized by a 10 bp periodicity of A/T as well as G/C sequence motives that are shifted by 5 bp relative to each other. As a consequence, this DNA has an intrinsic curvature that favors assembly into a nucleosome with A/T rich sequences strongly preferring minor groove sites that bind to the histone octamer and G/C rich sequences preferring major groove sites facing outwards. The advantage of the 601 sequence is that it binds histone octamers at a well defined position and therefore allows the preparation of homogeneous nucleosome samples. Homogeneity in the nucleosome position of an ensemble of nucleosomes prior to remodeling is crucial for being able to interpret smFRET remodeling data. However, the strong positioning sequence 601 also introduces artifacts since the 10 bp periodicity of strong histone-DNA interaction sites is likely to influence the mechanism of remodeling by forcing a step size of a multiple of 10 bp. Moreover, nucleosomes containing the 601 sequence are substantially more stable than nucleosomes on natural positioning sequences and hence the remodeler encounters an unnaturally high energy barrier during remodeling. Therefore, future experiments should test the remodeling mechanism presented here with different nucleosomal DNA sequences in order to define the influence of DNA sequence on the remodeling mechanism and to rule out possible artifacts of the 601 sequence. Alternatively, the energetic barriers of nucleosomes assembled on the 601 sequence could be weakened by using measurement buffers with a higher salt concentration.

Similar experiments as presented here for yeast Chd1 should be carried out with other

remodeling machines, starting from human Chd1 and other members of the CHD family up to members of other remodeler families such as ISWI or SWI/SNF. In this way, the remodeling characteristics of different remodelers could be compared in order to understand how different remodelers are able to achieve different remodeling outcomes.

Finally, in this work mononucleosomes were used as remodeling substrate even though remodelers never encounter single nucleosomes in the natural environment of a living cell. Especially for the spacing activity of Chd1 or ISWI remodelers, the presence of neighboring nucleosomes might be a crucial factor for the mechanism of remodeling. Hence, it remains to be shown how the simplistic mechanism obtained from the *in vitro* studies presented here has to be adapted to describe the complexity in a living cell. The first step towards a more realistic picture would be the analysis of Chd1 remodeling of di-nucleosome substrates, in which case the application of three-color FRET might be required. Moreover, the effect of Chd1 and other remodeling enzymes on the structure and dynamics of nucleosomal arrays as well as on higher order chromatin structures such as 30 nm chromatin fibers could be investigated by using labeled chromatin fibers and super-resolution fluorescence microscopy techniques such as STORM or STED [264].

In conclusion, we have only begun to exploit the tremendous potential of single molecule fluorescence techniques in uncovering molecular mechanisms of chromatin remodeling processes and future studies promise exciting new insights that will finally complete our understanding of this essential cellular process.

4 Single molecule FRET studies of nucleosome remodeling by Chd1

Table 4.6.1.: Buffers used in experimental procedures described in this section.

Buffer name	Composition	Application
SEC buffer	- 20 mM Tris/HCl pH 8.0 - 1 mM EDTA - 100 mM NaCl	SEC purification of nucleosomal DNA
5 x TBE buffer	- 446 mM Tris base pH 8.0 - 445 mM Boric acid - 13 mM EDTA	native PAA gelelectrophoresis
Histone unfolding buffer	- 20 mM Tris/HCl pH 7.5 - 7 M Guanidinium HCl - 10 mM DTT	Preparation of histone octamers
Histone refolding buffer	- 20 mM Tris/HCl pH 7.5 - 2 M NaCl - 1 mM EDTA - 5 mM β -Mercaptoethanol	Preparation of histone octamers
Chd1 lysis buffer	- 20 mM Tris/HCl pH 7.5 - 350 mM NaCl - 0.05 % β -Mercaptoethanol - 0.1 % Tween-20 - Protease inhibitors	Chd1 purification
Chd1 prescission buffer	- 20 mM Tris/HCl pH 7.5 - 350 mM NaCl - 1 mM DTT - 1 mM EDTA - 0.1 % Tween-20	Chd1 purification
Chd1 storage buffer	- 20 mM Tris/HCl pH 7.5 - 1 mM EDTA - 350 mM NaCl - 10 % Glycerol - 2 mM DTT or 0.5 mM TCEP	Chd1 storage at -80°C and labeling
No/Low/High salt buffer	- 10 mM Tris/HCl pH 7.5 - 1 mM EDTA - 0.05 % NP40 (Tergitol) - 1 mM β -Mercaptoethanol - 0 mM/ 50 mM/ 2M NaCl	Nucleosome assembly by salt gradient dialysis
Remodeling buffer	- 20 mM Tris/HCl pH 7.5 - 0.5 mM EGTA - 50 mM NaCl - 3 mM MgCl_2 - 10 % Glycerol - 2 mM DTT	Chd1 remodeling, binding and localization experiments
Oxygen Scavenger buffer	- 20 mM Tris/HCl pH 7.6 - 50 mM NaCl 40 % (w/v) D-(+)-glucose or 5 $\mu\text{g}/\mu\text{l}$ glucose oxidase or 217000 U/ml catalase	Glucose, glucose oxidase and catalase stocks

4.6. Experimental Procedures

4.6.1. Buffers

The composition and application of all buffers is listed in Table 4.6.1. Buffers were prepared using ultra-pure water (*Millipore Simplicity 185* or *Millipore Synergy*) and were sterile filtered and degassed prior to usage. Reducing agents DTT and β -Mercaptoethanol were added shortly before using the buffer.

Figure 4.6.1.: Schematic illustration of the nucleosomal DNA and designed label positions, depicted as stars (red, acceptor dye Alexa647 or Atto647N; yellow, donor dye Tamra or Atto532).

Mono-nucleosomes were assembled on 200 bp long DNA containing the strong nucleosome positioning sequence 601 [111]. The DNA sequence together with designed positions of dye labels is shown in Figure 4.6.1. Nucleotides marked as open circles constitute the linker DNA, whereas nucleotides shown as filled circles are part of the 601 sequence and therefore constitute DNA inside the nucleosome core. Hence, nucleosomes are assembled close to the end of the DNA template with a short 6 bp linker DNA on one side and a longer linker of 47 bp on the other side (*6-N-47*). The nucleosome dyad runs through the base-pair marked in red and defines the center of the nucleosome, around which it has a pseudo two-fold symmetry. All nucleotides and therefore all label positions are numbered relative to the dyad (= 0) with positive numbers towards the 3' end and negative numbers towards the 5' end of each strand. Accordingly, nucleotides on opposite strands of the DNA hold the same number but with opposite sign. The possible label positions are named according to this numbering. The reverse strand is modified with a biotin at the long linker side (5' end) to allow for immobilization to the quartz glass slide of the microfluidic measurement chambers.

Labeled nucleosomal DNA was prepared either by ligase chain reaction (LCR) or by polymerase chain reaction (PCR). I will introduce both preparation procedures in the

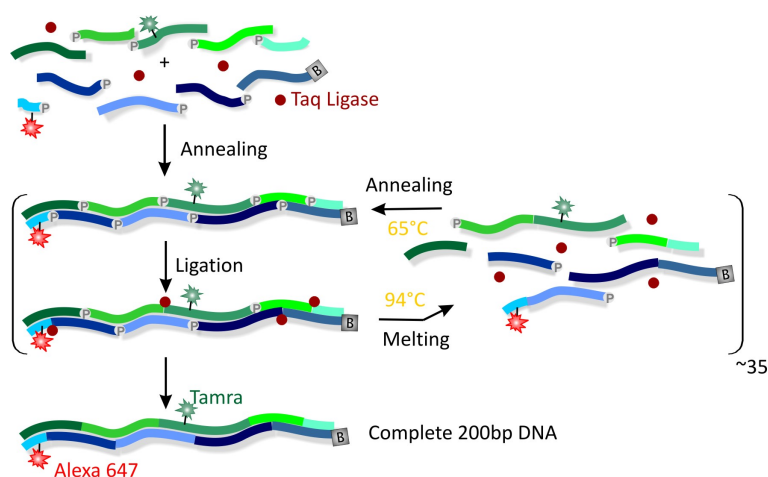


Figure 4.6.2.: Principle mechanism of the ligase chain reaction (LCR) catalyzed by Taq DNA ligase.

following and discuss their advantages and disadvantages.

4.6.3. LCR preparation of labeled nucleosomal DNA

In the LCR preparation, labeled 200 bp long DNA is built up from ten oligonucleotide fragments (IBA) (five fragments per strand) through ligation by the Taq DNA ligase (*New England BioLabs*). Oligonucleotide fragments were designed with the web-based tool *Gene2Oligo* [265], in which the length of the fragments is dynamically optimized to ensure both specificity and uniform melting temperatures. Table 4.6.2 shows all oligonucleotides in 5'→3' direction as well their respective length. Possible label positions are shown in bold and colored green for Atto532 or Tamra and red for Alexa647 or Atto647N, but also the unlabeled form of each oligonucleotide was available. As required for ligation by the Taq ligase, all fragments were phosphorylated at the 5' end except for the ones located at the 5' ends of both strands.

The main principle of the ligase chain reaction is shown in Figure 4.6.2. Equimolar amounts (typically 100 pmol) of all desired oligonucleotide fragments were mixed with Taq ligase (1 μ l of 40000 U/ml) in a final volume of 30 μ l Taq ligase buffer. In a thermal cycler equipped with a heated lid (99 °C), the mixture was heated to 94 °C for 2 min to denature the DNA. Then, 35 cycles of the following two-step program were performed: 6 min at 60 °C for hybridization of oligonucleotides and ligation of phosphorylated 5' ends to adjacent 3' ends by the Taq ligase followed by 30 s at 94 °C for denaturing of DNA. Finally, the mixture was slowly cooled down from 95 °C to 4 °C at a rate of 1 °C/min to ensure proper annealing. The LCR product was purified by using a PCR purification kit (*Qiagen*) followed by size exclusion chromatography (see below).

Table 4.6.2.: Oligonucleotides used for LCR preparation of 200 bp long nucleosomal DNA. The bold, colored bases represent the possible label positions (green, Atto532 or Tamra; red, Alexa647 or Atto647N). The name of each fragment is composed of either 'F' (forward strand) or 'R' (reverse strand) and a number that implies the order of ligation to yield the full-length product.

Name	label	Sequence (5'→3')	length /nt
F1	F(-64) Atto647N/ Alexa647	GGC CGC CCT GGA GAA TCC CGG TGC CGA GGC CGC TCA ATT G	40
F2	F(-15) Atto532/ Tamra	GTC GTA GCA AGC TCT AGC ACC GCT TAA ACG CAC GTA CGC GCT G	43
F3	F(+14) Atto532/ Tamra	TCC CCC GCG TTT TAA CCG CCA AGG GGA TTA CTC CCT AGT CTC CAG	45
F4	-	GCA CGT GTC AGA TAT ATA CAT CCT GTG CAT GTA TTG AAC AGC GAC TCG GGT TA	53
F5	-	TGT GAT GGA CCC TAT ACG C	19
R1	-	GGG ATT CTC CAG GGC GGC C	19
R2	-	CGG TGC TAG AGC TTG CTA CGA CCA ATT GAG CGG CCT CGG CAC C	43
R3	-	TTG GCG GTT AAA ACG CGG GGG ACA GCG CGT ACG TGC GTT TAA G	43
R4	R(-60) Atto532	GCA CAG GAT GTA TAT ATC TGA CAC GTG CCT GGA GAC TAG GGA GTA ATC CCC	51
R5-Biotin	R(-85) Alexa647/ Atto647N	Biotin- GCG TAT AGG GTC CAT CAC ATA ACC CGA GTC GCT GT CAA TAC AT	44

The advantage of the LCR preparation of labeled DNA lies in the flexibility with respect to label positions. Labels can be introduced at any position and even two labels far apart on the same final DNA strand are possible. However, I had to realize that, even at a high number of ligation cycles and high concentration of Taq ligase, the ligation reaction is never complete and thus, nicks remain in the DNA. The nicks present an obstacle for the remodeling enzyme Chd1 and therefore cause remodeling deficiencies. Therefore, PCR using labeled primers was introduced as a better preparation method of labeled nucleosomal DNA and most experiments presented in this thesis were performed with nucleosomes assembled from such PCR-based DNA. Nevertheless, PCR holds the disadvantage that the introduction of a fluorescent label is limited to the primers. Therefore, only a limited region of nucleosomal DNA is accessible for a label.

4.6.4. PCR preparation of labeled nucleosomal DNA

In the PCR preparation, labeled primers (*IBA*, see Table 4.6.3) were extended by the Phusion DNA polymerase (*Finnzymes*) to yield the full-length 200 bp DNA product. Typically, 200 ng of unlabeled template DNA (200 bp) was mixed with 50 pmol of each primer, 200 μ mol of each dNTP and Phusion DNA polymerase (final concentration 0.05 U/ μ l) in a final volume of 50 μ l Phusion buffer. In a thermal cycler equipped with a heated lid (heated to 99 °C), the mixture was heated to 95 °C for 5 min to denature the DNA.

4 Single molecule FRET studies of nucleosome remodeling by Chd1

Table 4.6.3.: List of DNA primers used for PCR preparation of nucleosomal DNA. The primers are named according to the label position (bold colored base in sequence), the dye and the sense (F, forward primer; R, reverse primer).

Name	Sequence (5'→3')	length/nt
F non-labeled	GGC CGC CCT GGA GAA TCC CGG TGC CGA GGC CGC TCA ATT GGT CGT AGC AAG CTC TAG CAC CGC TTA AAC GCA CGT ACG CGC TTA AAC GCA CGT ACG CGC TGT CCC CCG CGT TTT AAC C	100
F(-term)-Alexa647	ALX -GGC CGC CCT GGA GAA TCC CGG TGC CGA GGC CGC TCA ATT GGT CGT AGC AAG CTC TAG CAC CGC TTA AAC GCA CGA CGT ACG CGC TTA AAC GCA CGT ACG CGC TGT CCC CCG CGT TTT AAC C	100
F(+14)-Tamra/Atto532	GGC CGC CCT GGA GAA TCC CGG TGC CGA GGC CGC TCA ATT GGT CGT AGC AAG CTC TAG CAC CGC TTA AAC GCA CGT ACG CGC TTA AAC GCA CGT ACG CGC TGT CCC CCG CGT T TT AAC C	100
F(+2)-Tamra	GGC CGC CCT GGA GAA TCC CGG TGC CGA GGC CGC TCA ATT GGT CGT AGC AAG CTC TAG CAC CGC TTA AAC GCA CGT ACG CGC T GT CCC CCG	90
F(-15)-Tamra/Atto532	GGC CGC CCT GGA GAA TCC CGG TGC CGA GGC CGC TCA ATT GGT CGT AGC AAG CTC TAG CAC CGC T TA AAC GCA C	73
F(-64)-Tamra	GGC CGC CCT GGA GAA T CC CGG TGC CGA GGC CGC TCA ATT GGT CGT AGC AAG CTC TAG CAC CGC TTA AAC GCA C	73
F(+14)-Tamra-F(-64)-Alexa647	GGC CGC CCT GGA GAA T CC CGG TGC CGA GGC CGC TCA ATT GGT CGT AGC AAG CTC TAG CAC CGC TTA AAC GCA CGT ACG CGC TTA AAC GCA CGT ACG CGC TGT CCC CCG CGT T TT AAC C	100
F(+14)-Tamra-F(-71)-Alexa647	GGC CGC C T GGA GAA TCC CGG TGC CGA GGC CGC TCA ATT GGT CGT AGC AAG CTC TAG CAC CGC TTA AAC GCA CGT ACG CGC TTA AAC GCA CGT ACG CGC TGT CCC CCG CGT T TT AAC C	100
R non-labeled	Biotin- GCG TAT AGG GTC CAT CAC ATA ACC CGA GTC GCT GTT CAA TAC	42
R(-60)-Tamra	Biotin- GCG TAT AGG GTC CAT CAC ATA ACC CGA GTC GCT GTT CAA TAC ATG CAC AGG ATG TAT ATA T CT GAC ACG	69
R(-66)-Tamra	Biotin- GCG TAT AGG GTC CAT CAC ATA ACC CGA GTC GCT GTT CAA TAC ATG CAC AGG ATG T AT ATA TCT GAC ACG	69
R(-85)-Alexa647/Tamra	Biotin- GCG TAT AGG GTC CAT CAC ATA ACC CGA GTC GCT GT T / T CAA TAC	42

Then, the DNA was amplified in 35 repetitions of the following PCR cycle: 30 s denaturation of DNA at 95 °C, 1 min primer annealing at 68 – 72 °C (lower temperature for shorter primers) and 1 min primer extension by the DNA polymerase at 72 °C. The cycle was followed by a final polymerization step at 68 – 72 °C (lower temperature for shorter primers) for 10 min. The PCR product was purified using a PCR purification kit (*Qiagen*) followed by size exclusion chromatography (SEC) (see below).

4.6.5. SEC purification of labeled nucleosomal DNA

Nucleosomal DNA prepared by LCR or PCR was subjected to a final SEC purification step. To this end, 50 to 100 μ l of the crude DNA product was loaded onto a Superose 6 PC 3.2/30 column (*GE Healthcare*) equilibrated in SEC buffer (Table 4.6.1) and eluted at a flow rate of 0.020 ml/min. The peak fractions were analyzed on a 6 % PAA gel (1xTBE, 20 min at 300 V). Fluorescent gel bands were visualized using a Typhoon Imager (*GE*

Healthcare) (Tamra, Atto532: excitation at 532 nm, detection with emission filter 580 BP 30; Alexa647, Atto647N: excitation at 647 nm, detection with emission filter 670 BP 30) and subsequent to SYBR[®] Gold staining, the total DNA content of the gel was imaged on a Gelimager (*Biorad*) (excitation: 300nm, detection: ~570 nm). Fractions containing the desired DNA were pooled and concentrated to about 20 μ l using Amicon[®] Ultra30K concentrators (*Millipore*). The DNA concentration was determined via absorption at 260 nm using a NanoDrop[®] ND-1000 (*Peqlab*) (after removal of residual dye signal) and the DNA was stored at -20°C .

4.6.6. Unlabeled nucleosomal and competitor DNA

For assembly of fluorescently labeled nucleosomes, not only labeled nucleosomal DNA, but also unlabeled nucleosomal DNA as well as long competitor DNA is required. Moreover, unlabeled nucleosomal DNA is needed as template in the PCR preparation of labeled DNA.

Unlabeled 200 bp long DNA and competitor DNA were prepared from a pUC18 plasmid containing an insert of twelve repeats of 200 bp nucleosomal DNA (Figure 4.6.1) separated by NotI restriction sites (pUC18-12x601, gift from G. Längst). NotI digestion of the plasmid yielded twelve fragments of 200 bp nucleosomal DNA as well as the long backbone DNA (pUC18 vector) used as competitor DNA.

Prior to digestion, the pUC18-12x601 plasmid was amplified in *Escherichia Coli* XL1 blue competent cells (*Stratagene*). Thereto, competent cells were thawed on ice, plasmid DNA was added (1 ng plasmid-DNA per 100 μ l competent cells) to the cells on ice and incubated for 2 min. Subsequently, cells were exposed to a heat shock at 42°C for 45 s followed by incubation on ice for 10 min. 750 μ l LB-medium (1 % tryptone, 0.5 % yeast extract, 0.5 % NaCl, pH 7.0) was added and the cells were incubated at 37°C for 1 h. After centrifugation (4000 rpm, 2 min, RT) and removal of the supernatant, the cell pellet was resuspended in the remaining LB-medium and plated on LB-agar plates containing ampicillin (*ROTH*). The agar-plates were incubated over night at 37°C . On the next day, 100 ml LB-medium containing 100 μ l ampicillin was inoculated with a colony of grown bacteria and shaken at 37°C over night. Cells were pelleted by centrifugation (4000 rpm, 10 min, RT) and plasmid DNA was extracted using a Miniprep Kit (*Qiagen*).

The pUC18-12x601 plasmid was digested by NotI-HFTM (*New England Biolabs*) for 16 h at 37°C followed by a heat inactivation step at 65°C for 20 min. Subsequently, the DNA was purified using a PCR Purification Kit (*Qiagen*). For nucleosome assembly, purified digestion product containing 200 bp nucleosomal DNA as well as long backbone DNA (*competitor*) was directly used. However, in order to obtain template DNA for PCR preparation of labeled nucleosomal DNA, the 200 bp nucleosomal DNA was separated from backbone DNA on a 1 % agarose gel (40 min, 100 V), excised from the gel and extracted using a Gel Extraction Kit (*Metabion*). In both cases, DNA concentration was determined via absorption at 260 nm using a NanoDrop[®] ND-1000 (*Peqlab*) prior to

storage at -20°C .

4.6.7. Preparation of Histone Octamers

Initially, mono-nucleosomes were assembled from endogenous chicken histone octamers, provided by the laboratory of Prof. Gernot Längst (University of Regensburg). Octamers were purified from chicken erythrocytes, thus bearing all naturally occurring post-translational modifications. From those histones, only a heterogeneous population of nucleosomes with uncontrollable chemical modifications of histones could be prepared. Therefore, later on in this study, endogenous chicken histone octamers were replaced by recombinant human histone octamers, which are completely free of post-translational modifications, thus yielding homogenous nucleosome samples. Most data presented in this thesis was acquired with nucleosomes containing recombinant human histones. Recombinant human histone octamers were prepared in our laboratory by Katarzyna Krzemień and Monika Holzner following a protocol developed in the laboratory of Prof. Peter Becker (Adolf-Butenandt Institute, Ludwig-Maximilians-Universität München) by Felix Müller-Planitz (unpublished results). All four histones were separately expressed in *Escherichia Coli* BL21(DE3) cells and purified using tandem cation- and anion-exchange chromatography. Finally, histones were lyophilized and stored at -80°C .

Reconstitution and purification of histone octamers was performed as described in [266] with slight modifications. Briefly, each histone was unfolded in Histone unfolding buffer (Table 4.6.1) for 0.5-3 h at room temperature and the concentration of each histone was determined via absorption at 276 nm ($\epsilon(H2A)_{276\text{nm}} = 4050 \text{ M}^{-1}\text{cm}^{-1}$, $\epsilon(H2B)_{276\text{nm}} = 6070 \text{ M}^{-1}\text{cm}^{-1}$, $\epsilon(H3)_{276\text{nm}} = 4040 \text{ M}^{-1}\text{cm}^{-1}$, $\epsilon(H4)_{276\text{nm}} = 5400 \text{ M}^{-1}\text{cm}^{-1}$). The four histone proteins were mixed to equimolar ratios at a final concentration of 1 mg/ml and refolded by dialysis against three times 2 l of Histone refolding buffer (Table 4.6.1), twice for 2 h and the third time overnight. After centrifugation, histone octamers were purified from aggregates, tetramers and dimers by size exclusion chromatography using a Superdex 200 column (*GE Healthcare*) equilibrated in Histone refolding buffer at a flow rate of 1 ml/min. Octamers usually eluted at a retention volume of around 65-68 ml, tetramers and dimers eluted at a larger volume. Fractions containing histone octamers were pooled, concentrated to about 20 μl using Amicon[®] MWCO 10 K concentrators (*Millipore*) and glycerol was added to a total concentration of 50 %. Finally, histone octamers were stored at -20°C for several months.

4.6.8. Nucleosome assembly via salt gradient dialysis

Nucleosomes, particularly at picomolar dilutions and/or in the presence of salt, are prone to disassemble mainly due to non-specific binding of histones to sticky surfaces [115, 267, 268]. Therefore, siliconized microcentrifuge tubes as well as siliconized pipette tips were used whenever nucleosomes were contained in the handled solution.

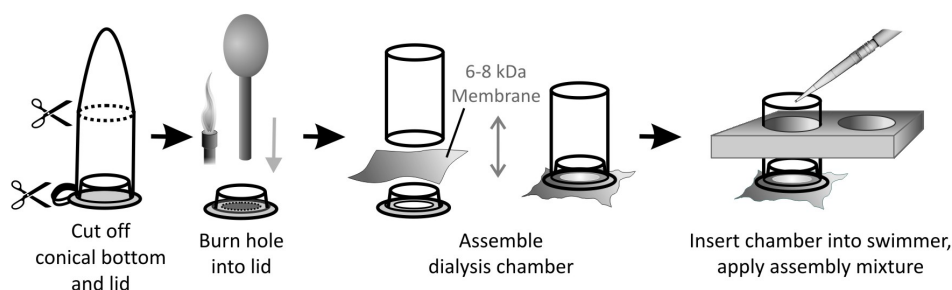


Figure 4.6.3.: Schematic illustration of the preparation of dialysis chambers for nucleosome assembly via salt gradient dialysis.

End-positioned mono-nucleosomes were reconstituted from DNA and histones by salt gradient dialysis as described by Rhodes [269]. Histones are stable as monomers in high salt conditions, but specifically assemble onto DNA during dialysis to low salt conditions- first the central (H3-H4)₂ tetramer, then the peripheral H2A-H2B dimers. Labeled nucleosomal DNA and unlabeled DNA (digestion product of pUC18-12x601 plasmid containing nucleosomal and competitor DNA, see section 4.6.6) were used in a mass ratio of 1:19. Different mass ratios of histones to DNA (usually in the range of 0.5:1 to 1.5:1) were tested to find the optimal conditions ensuring complete compaction of nucleosomal DNA. The long, unlabeled competitor DNA served as a kind of buffer by non-specifically binding excess histones and thus preventing assembly of non-canonical nucleosomes. Nucleosome assembly was performed in small dialysis chambers built from siliconized 1.5 ml reaction tubes (*Biozym*) as outlined in Figure 4.6.3. Conical bottom and lid of a siliconized tube were removed using scissors. A hole was melted into the lid using a heated metal rod (spoon-handle or spatula) and any remaining sharp plastic protrusions were removed using a scalpel. A piece of 6.0-8.0 kDa dialysis membrane (*spectrapor*) was pre-equilibrated in High salt buffer (Table 4.6.1) for 10 min and then clipped as a single layer between the perforated lid and the remaining cylindrical top of the tube. The resulting dialysis chamber was inserted into a styrofoam swimmer and placed into a 5 l bucket filled with 300 ml High salt buffer. Air bubbles below the membrane were removed using a bent Pasteur pipette.

Finally, the assembly reaction mixture was applied to the dialysis chamber, which typically consisted of 5 μ g DNA (250 ng labeled, 4.75 μ g unlabeled) and varying amounts of histones (2.5-7.5 μ g) in a total volume of 40 μ l of High salt buffer supplemented with 200 ng/ μ l BSA. Salt gradient dialysis was performed at 4 °C by continuously adding Low salt buffer (Table 4.6.1) to the bucket containing the dialysis chambers in High salt buffer by using a peristaltic pump. A total volume of 3 l Low salt buffer was added to the 300 ml of High salt buffer at a flow rate of 200 ml/h and hence, over a period of 15 h, the NaCl concentration dropped from 2 M to 227 mM. The assembly was completed by a final dialysis step against 2 l of No salt buffer (Table 4.6.1) for 2 h. The concentration of the resulting nucleosomes

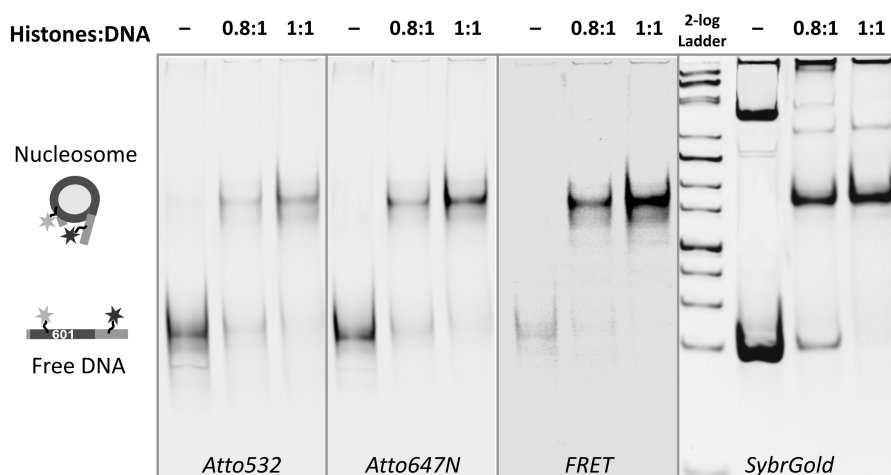


Figure 4.6.4.: Exemplary electrophoretic mobility shift assay demonstrating nucleosome assembly via salt gradient dialysis. Nucleosomes assembled from double-labeled nucleosomal DNA (R(-85)-Atto647N and F(+14)-Atto532, LCR-based) and chicken histones were assayed on a 6 % PAA gel (0.4 x TBE). Green fluorescence of Atto532 (excitation: 532 nm, detection: 565-595 nm), red fluorescence of Atto647N (excitation: 633 nm, detection: 655-685 nm) as well as the FRET signal (excitation: 532nm, detection: 655-685 nm) between the two dyes are displayed. Moreover, the signal of total DNA upon SYBR[®] Gold staining is shown (excitation: 300nm, detection: ~570 nm). Binding of 200 bp nucleosomal DNA to histones during nucleosome assembly results in an upshift of the DNA signal in the gel. Two different histone to DNA ratios were tested (0.8:1 and 1:1) and the ratio of 1:1 proved optimal as all nucleosomal DNA is compacted. Whereas free double-labeled DNA does not exhibit FRET, a notable FRET signal could be detected from the double-labeled DNA within nucleosomes. In the SYBR[®] Gold stained gel, long competitor DNA is visible that also shows an upshift due to histone binding. Moreover, no preferential binding of histones to either labeled or unlabeled DNA could be observed.

was estimated from the change in volume during the dialysis and was expressed as DNA mass per volume (typically in the range of 90-125 ng/ μ l). Nucleosomes were assayed on a 6 % PAA gel (0.4 x TBE, 90 min, 100 V) (Figure 4.6.4). Typically, per assembly reaction 400 ng referring to DNA mass was loaded in water supplemented with 2.5 % Ficoll (*Sigma-Aldrich*) to facilitate loading. Fluorescent bands were visualized using a Typhoon Imager (*GE Healthcare*) followed by SYBR[®] Gold staining and visualization of the total DNA on a Gelimager (*Biorad*). Finally, nucleosomes were stored at 4 °C in siliconized tubes.

4.6.9. Expression and Purification of Chd1

All constructs of *Saccharomyces cerevisiae* Chd1 were provided by Daniel Ryan from the laboratory of Prof. Tom Owen-Hughes (College of life sciences, University of Dundee). In all experiments, a truncated version of Chd1 was used lacking the C-terminus (Chd1 1-1305 Δ C) (Figure 4.6.5). The C-terminal truncation was previously shown to have no affect on nucleosome binding, ATPase activity or nucleosome repositioning [149]. In remodeling

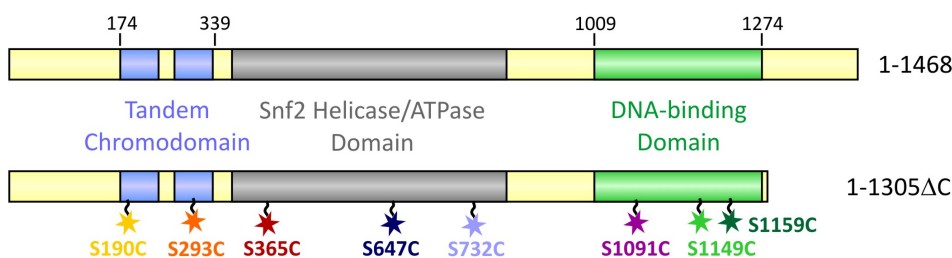


Figure 4.6.5.: Domain structure of Chd1. The full-length protein as well as the used truncated version lacking the C-terminus (1-1305ΔC) are shown.

experiments, wild-type Chd1 1-1305ΔC was used, whereas in NPS localization experiments Chd1 1-1305ΔC mutants were used lacking all native cysteines (C207S, C246S, C453S, C455S, C1063S) and having one of the following serines mutated to a cysteine by site-directed mutagenesis: S190C, S293C, S365C, S647C, S732C, S1091C, S1149C or S1159C (Figure 4.6.5). Additionally, one single cysteine construct was lacking all native cysteines but C1063 (C207S, C246S, C453S, C455S). Wild-type Chd1 1-1305ΔC as well as single cysteine mutants were expressed and purified as described previously [149, 270]. Briefly, a pGEX-6P expression vector (*GE Healthcare*) from yeast genomic DNA containing the coding sequence for full-length wild-type Chd1 with a C-terminal 6 x His-tag was modified by inverse PCR strategy to yield the C-terminal truncation. Point mutations were introduced by site-directed mutagenesis using a QuikChange[®] kit (*Agilent*). Wild-type as well as mutant Chd1 1-1305ΔC was expressed in *Escherichia Coli* Rosetta 2 (DE3) pLysS cells at 20 °C for 48-60 h in autoinduction media containing ampicillin and chloramphenicol (both *ROTH*). Cells were harvested by centrifugation (5000 g, 12 min, 4 °C) and resuspended in Chd1 lysis buffer (Table 4.6.1) containing protease inhibitors (0.2 mM AEBSF, 2.6 mM aprotinin, 1 μM pepstatin, 2 μM E64). Cells were lysed by freeze-thawing in liquid nitrogen followed by sonication and the soluble fraction containing GST-Chd1 1-1305ΔC-6xHis was isolated by centrifugation for 30 min at 35 000 g and 4 °C. Chd1 was purified by means of tandem glutathione- and metal-affinity purification as described similarly for the Fun30 chromatin-remodeling enzyme [270]. Briefly, GST-Chd1 1-1305ΔC-6xHis was applied to SuperGlu glutathione resin (*Generon*) and washed consecutively with Chd1 lysis buffer (Table 4.6.1) and Chd1 prescission buffer (Table 4.6.1). Bound protein was released from GST by cleavage with Prescission Protease (home-made) overnight at 4 °C and eluted in Chd1 prescission buffer lacking EDTA and DTT. The eluate containing Chd1 1-1305ΔC-6xHis was further applied to HisPur cobalt resin (*Thermo Scientific*), washed with Chd1 lysis buffer and eluted from the resin using 250 mM imidazole in Chd1 lysis buffer. Chd1 1-1305ΔC-6xHis was then concentrated to a volume of 100-200 μl in 50 KDa MWCO centrifugal concentrators (*Millipore*), dialyzed against Chd1 storage buffer (Table 4.6.1) and finally stored at −80 °C.

4.6.10. Labeling of Chd1

Single cysteine mutants of Chd1 were labeled with the thiol-reactive dye derivative Alexa647-C2-Maleimide (*Molecular Probes*). Chd1 was kept on ice or in the fridge at 4 °C throughout the whole labeling procedure as exposure to room temperature was found to harm the remodeling activity of the enzyme. First, an aliquot (usually 30 μ l) of Chd1 with the desired single cysteine mutation was thawed, placed on ice and supplied with fresh DTT to a final concentration of 2 mM in order to fully reduce the protein. After 10 min incubation, the DTT bearing buffer was exchanged against Chd1 storage buffer (Table 4.6.1) without DTT using G-50 spin columns (*GE Healthcare*) and the protein concentration (usually around 2 μ M) was determined via absorption at 280 nm using a NanoDrop[®] ND-1000 (*Peqlab*) ($\epsilon(\text{Chd1})_{280\text{nm}} = 143840 \text{ M}^{-1}\text{cm}^{-1}$). Meanwhile, Alexa647-C2-Maleimide was dissolved in 5 μ l of water and its concentration was determined via absorption at 647 nm ($\epsilon(\text{Alexa647})_{649\text{nm}} = 239\,000 \text{ M}^{-1}\text{cm}^{-1}$). Subsequently, a 20-fold molar excess of Alexa647-C2-Maleimide was added to the Chd1 solution and the labeling reaction was allowed to occur over night at 4 °C. To stop labeling, the reaction mixture was supplied with DTT at a final concentration of 2 mM. Subsequently, unreacted dye was removed using G-50 spin columns (*GE Healthcare*) with Remodeling buffer (Table 4.6.1) containing 2 mM DTT. Typically, two to three consecutive spin-column purification steps were necessary to fully remove free dye. Finally, concentration and labeling efficiency of Chd1 were determined via absorption at 280 nm and 647 nm.

4.6.11. Chd1 Binding assay

Binding of Chd1 to mono-nucleosomes was investigated by an electrophoretic mobility shift assay. Typically, in a total volume of 12 μ l Remodeling buffer (Table 4.6.1) containing 2.5% Ficoll (*Sigma-Aldrich*), 0.5 μ l of donor labeled mono-nucleosomes (approximately 30 nmol) were incubated with varying amounts of Alexa647 labeled Chd1 in the presence of 1 mM of the non-hydrolyzable ATP analog AMP-PNP on ice. Different Nucleosome to Chd1 ratios were tested to find the optimal binding conditions, usually ranging from 1:1 (30 nmol Chd1) to 1:15 (450 nmol Chd1). After incubation for 10 min, the different binding reactions were loaded onto a 6 % PAA gel (0.4 x TBE) (prerun at 4 °C, 60V for 30 min) after thoroughly rinsing the pockets. The gel was run for 90 min at 100 V and at 4 °C. Fluorescent bands were visualized using a Typhoon Imager (*GE Healthcare*) followed by SYBR[®] Gold staining and visualization of the total DNA on a Gelimager (*Biorad*). In a binding assay performed in the early stages of this project, different ATP analogs were tested for their stimulating effect on the binding of Chd1 to mono-nucleosomes (Figure 4.6.6). Compared to ATP- γ S and ADP-BeF_x, AMP-PNP most strongly promoted Chd1 binding and was therefore used in all subsequent Chd1 binding and positioning experiments.

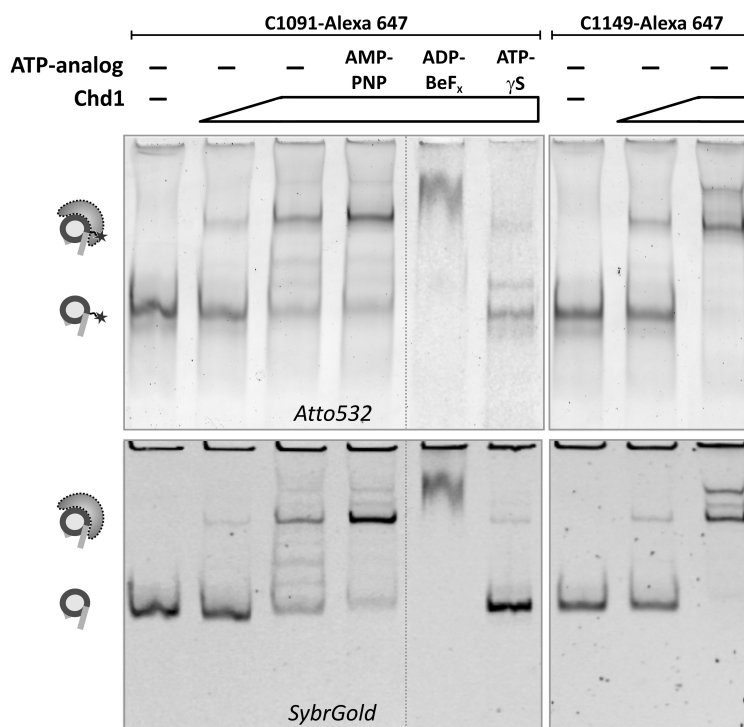


Figure 4.6.6.: Binding of Chd1-C1091-Alexa647 (left) and Chd1-C1149-Alexa647 (right) to labeled mono-nucleosomes F(-15)-Atto532. The green fluorescence of Atto532 (upper panel) as well as the signal of the total DNA upon SYBR[®] Gold staining (lower panel) are shown. An upshift of the nucleosome band is observed upon Chd1 binding. Two different nucleosome to Chd1 ratios were tested (1:2 and 1:15) as well as the effect of different non-hydrolyzable ATP analogs on the binding activity in the case of Chd1-C1091 (AMP-PNP, ATP- γ S and ADP-BeF_x). AMP-PNP appears to promote binding of Chd1 to mono-nucleosomes most prominently. Moreover, Chd1 does not seem to discriminate between labeled and unlabeled nucleosomes regarding binding. For the higher concentration of Chd1-C1149-Alexa647 (rightmost lane), a supershift can be observed indicating binding of more than one Chd1 remodeler per nucleosome.

4.6.12. Nucleosome remodeling assay

Chd1 preferentially moves end-positioned nucleosomes, like the ones used in this work (6-N-47), towards a more central location on a short DNA fragment in order to achieve similar linker lengths on both sides of the nucleosome [128]. The fact that end-positioned nucleosomes migrate faster than centered nucleosomes in a native gel [271] is harnessed in the remodeling assay, because successful remodeling can be detected as an upshift of the band of an initially end-positioned nucleosome. Nucleosome remodeling was carried out in a 12 μ l reaction containing 3 pmol of fluorescently labeled nucleosomes, typically 1.2 pmol of Chd1 (either wild-type or labeled mutant 1-1305 Δ C; omitted in negative control) and 150 μ M ATP in Remodeling buffer (Table 4.6.1). After 90 min incubation at 27 $^{\circ}$ C, the remodeling reaction was stopped by transfer onto ice and addition of 1 μ g of competitor

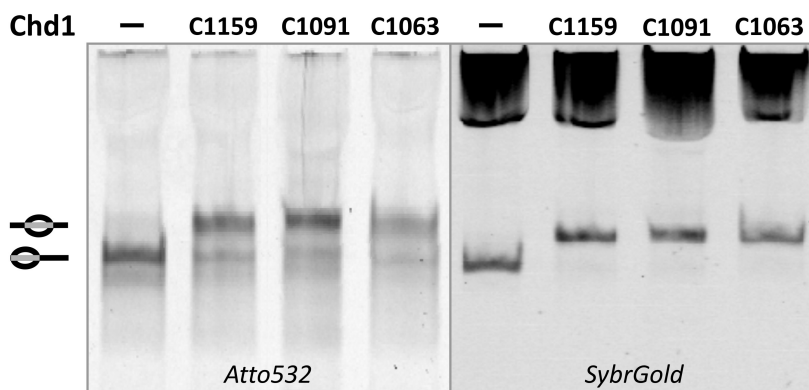


Figure 4.6.7.: Native gel analysis of nucleosome remodeling catalyzed by Alexa647-labeled Chd1 mutants Chd1-C1091, Chd1-C1063 and Chd1-C1159. Labeled Chd1 mutants are capable of sliding the end-positioned nucleosomes towards the center of the nucleosomal DNA F(-15)-Atto532 (LCR-based) apparent by an upshift (lane 2-4) of the initial nucleosome band (lane 1). The upper band in the SYBR[®] Gold stain image represents the unlabeled competitor DNA added to the reactions to stop the remodeling process by withdrawing Chd1 from the nucleosomal template.

DNA (bacteriophage -DNA (*New England BioLabs*). Finally, the remodeling product was analyzed on a 0.4 x TBE 6 % PAA gel (100 V, 90 min) followed by visualization using a Typhoon Imager (*GE Healthcare*) and upon SYBR[®] Gold staining using a Gelimager (*Biorad*). Figure 4.6.7 shows an exemplary remodeling assay, where successful remodeling by the Alexa647 labeled Chd1 mutants Chd1-C1091, Chd1-C1063 and Chd1-C1159 was demonstrated. Occasionally, the product of a remodeling assay as well as its negative control were analyzed in a smFRET experiment in addition to the native gel analysis. To this end, 0.5 μ l was subtracted from each desired remodeling reaction, diluted 100-1000 fold and immobilized in a microfluidic measurement chamber for further smFRET measurements.

4.6.13. Oxygen-scavenging system and Trolox

To follow conformational changes of a nucleosome during remodeling by Chd1 by means of smFRET, fluorophores attached to the nucleosome should not undergo fast photo-oxidation (bleaching) or transit to dark triplet states (photo-blinking) during an extended period of time. Long-lasting, nonblinking fluorophores can be achieved by using a measurement buffer supplemented with an enzymatic oxygen-scavenging system [272] in combination with the vitamin E analog Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid) [273]. The oxygen-scavenging system is composed of glucose oxidase and catalase, which use D-(+)-glucose and dissolved O₂ to build gluconic acid and water. This enzymatic reaction results in a net loss of molecular oxygen O₂ from solution and thus reduces bleaching of fluorophores by photo-oxidation. However, beside being responsible for

photo-bleaching, O_2 acts as triplet-state quencher and removal of O_2 from solution therefore leads to increased photoblinking. This negative effect is compensated by the addition of Trolox as an alternative triplet-state quencher. Trolox together with its oxidized form Trolox quinone, which forms during illumination of Trolox by UV-light, serves as oxidizing and reducing system (ROXS) that depopulates triplet-states through electron transfer [274]. Remodeling buffer (Table 4.6.1) supplemented with oxygen-scavenging system and Trolox (referred to as *OxScavTX-Remodeling buffer* in the following) was prepared as follows: 250 μ l of Trolox (*Fluka*, 2mM in Remodeling buffer, filtered, stored in the dark) was illuminated with UV-light for 6 min in order to create the proper equilibrium between Trolox and Trolox quinone. Meanwhile, the oxygen-scavenging system was prepared by mixing 2 μ l of glucose oxidase (*Sigma-Aldrich*; 5 μ g/ μ l in Oxygen Scavenger buffer, filtered), 1 μ l of catalase (*Roche*; 217000 U/ml in Oxygen Scavenger buffer, filtered) and 10 μ l of D-(+)-glucose (*Sigma-Aldrich*, 40 % (w/v) in freshly degassed Oxygen Scavenger buffer, filtered) in 87 μ l of Remodeling buffer containing 200 ng/ μ l BSA (100 μ l total volume). Finally, the oxygen-scavenger solution was combined with 100 μ l of the Trolox solution, allowed to equilibrate for 10 min and immediately used in the smFRET experiment. The anti-bleaching and anti-blinking effect of the buffer was found to be expended after approximately 30 min and hence, the buffer had to be renewed in experiments longer than that.

4.6.14. smFRET analysis of nucleosome dynamics induced by Chd1

Nucleosome dynamics induced by Chd1 through ATP-driven repositioning or through binding (in the presence or absence of AMP-PNP) were examined in time-resolved smFRET experiments using our home-build prism-based TIRF microscope (section 2.2.2). The experiments were performed using nucleosomes with donor-acceptor-labeled nucleosomal DNA (F(+14)-Tamra/ R(-85)-Alexa647, F(+14)-Tamra/ F(-71)-Alexa647 or R(-85)-Tamra/ F(-term)-Alexa647) as well as unlabeled Chd1 1-1305 Δ C. Repositioning experiments were conducted in the following way: A microfluidic measurement chamber was prepared as described in section 2.2.2. Upon coating the chamber surface with Neutravidin, it was further passivated by incubation with 1 mg/ml BSA in Remodeling buffer (Table 4.6.1) for 10 min followed by rinsing with Remodeling buffer containing 0.2 mg/ml of BSA. After pre-bleaching the chamber to reduce background fluorescence, nucleosomes (100-125 ng/ μ l referring to DNA mass) were diluted 5000-20 000 fold in Remodeling buffer (Table 4.6.1) and loaded into the measurement chamber to be immobilized onto the chamber surface via PEG-Biotin-Neutravidin-Biotin attachment. Subsequently, the chamber was rinsed with 200 μ l of OxScavTX-Remodeling buffer (section 4.6.13) containing 150 μ M ATP (*Sigma-Aldrich*) and 10-15 movies were recorded of nucleosomes under the control conditions without Chd1 but with ATP. Subsequently, Chd1 was loaded to the nucleosomes at a concentration of 50 nM in 100 μ l of freshly prepared OxScavTX-Remodeling buffer containing 150 μ M ATP and 15-25 movies were recorded of the nucleosomes in the

presence of Chd1 and ATP. All measurements were performed using alternating laser excitation (ALEX, alternation between excitation at 532 nm and at 637 nm) (section 2.2.1) at a frequency of 10 or 30 Hz and tif-movies were recorded with an integration time of 100 or 33 ms per frame, respectively, and for a total duration of 1000-2000 frames. Binding experiments were carried out similarly to the above described remodeling experiments, except that 150 μ M AMP-PNP (*Sigma-Aldrich* or *Roche*) or no ATP analog was used instead of ATP. In some cases, after having allowed Chd1 to bind nucleosomes in the presence or absence of AMP-PNP for about 10 min, unbound Chd1 was flushed out of the chamber using 400 μ l of freshly prepared OxScavTX-Remodeling buffer with or without AMP-PNP, respectively. Subsequently, OxScavTX-Remodeling buffer with 150 μ M ATP was loaded into the chamber and smFRET measurements were recorded to test whether Chd1 remained bound to nucleosomes and showed remodeling activity.

4.6.15. Hidden Markov Modeling (HMM) analysis of nucleosome dynamics

Dynamic smFRET time traces were selected and analyzed with the software *vbFRET* (variational bayesian inference for smFRET time series) [275] to identify the most probable number of states present in each dynamic time trace. The resulting number of states (four or five states for measurements with Chd1 and ATP and 2 states for measurements with Chd1 and AMP-PNP or Chd1 alone) was used as start parameter in the subsequent HMM analysis. HMM analysis was performed using a custom-written MATLAB (*The MathWorks*) software by Gregor Heiss and Martin Sikor based on the MATLAB HMM-Toolbox by Kevin Murphy [214] and on an algorithm similar to one previously described for HMM analysis of time-binned smFRET trajectories [212] (section 2.4). A five-, four- or two-state hidden Markov process was assumed and the width of states (sigma) was chosen to be 0.02 (2 % FRET) and was fixed during the analysis, since this setting yielded the best results with respect to the identification of transitions (see Supplemental Figures A.2.1 and A.2.2 and corresponding discussion). All transitions between states identified in the time-binned smFRET time trajectories by global HMM analysis were compiled in a transition density plot (TDP) (initial versus final FRET efficiency of every transition). In this process each transition was superimposed by a 2D Gaussian with a fixed width (due to fixed sigma during HMM analysis) into the TDP. The width was set to 0.04 (two times σ) for a better visualization compared to 0.02. The different transition populations in the TDP were selected and the cumulative events with a dwell time longer than a given time t were plotted against the time for each population. A single exponential-decay function was fitted to the cumulative distribution and the rate k of the transition could be extracted from the exponent. The standard deviation of the rate was obtained by building cumulative distributions from 10 different random subsets of one third of all transitions and fitting each single one with a mono-exponential decay function. In addition, from the relative occurrence of each transition, the transition probabilities could be calculated and were normalized such that the sum of probabilities for transitions

originating from one state was 1.

4.6.16. Chd1 positioning experiments

In order to position Chd1 relative to the bound nucleosome, smFRET experiments were performed between Tamra as satellite attached to one of six known positions on the nucleosomal DNA (F(+14), F(+2), F(-15), F(-64), R(-60) or R(-66)) and Alexa647 as antenna attached to one of eight positions on Chd1 (C190, C293, C365, C647, C732, C1091, C1149 or C1159). A microfluidic measurement chamber was prepared as described in section 2.2.2. Upon coating the chamber surface with Neutravidin, it was passivated by incubation with 1 mg/ml BSA in Remodeling buffer (Table 4.6.1) for 10 min followed by rinsing with Remodeling buffer containing 0.2 mg/ml of BSA. Donor-labeled Nucleosomes (100-125 ng/ μ l referring to DNA mass) were diluted 5000-20 000 fold in Remodeling buffer (Table 4.6.1) supplemented with 0.2 mg/ml BSA. After pre-bleaching the measurement to reduce background fluorescence, nucleosomes were loaded into the measurement chamber to be immobilized onto the chamber surface via PEG-Biotin-Neutravidin-Biotin attachment. After short incubation, unbound nucleosomes were flushed out of the chamber using Remodeling buffer (Table 4.6.1) containing 0.2 mg/ml BSA. Acceptor-labeled Chd1 was diluted to a final concentration of 50 nM in 100 μ l of Remodeling buffer containing 0.2 mg/ml BSA and 150 μ M AMP-PNP and centrifuged at 14 000 rpm for 10 min at 4 °C to separate monomers from aggregates. The upper 70 μ l were extracted and loaded into the measurement chamber. Chd1 was allowed to bind to immobilized nucleosomes for 5-10 min before unbound Chd1 was rinsed out of the chamber with 100-200 μ l of Remodeling buffer containing 0.2 mg/ml BSA and 150 μ M AMP-PNP. Sif movies were usually recorded with an integration time of 100 ms per frame and for a total duration of 40 s.

4.6.17. Global NPS analysis of a Chd1-nucleosome complex

Dye position priors

Satellite dye molecules were attached via flexible linkers to positions on the nucleosomal DNA consisting of the strong nucleosome positioning sequence 601 [111]. Due to the high stability of nucleosomes assembled onto the 601 sequence [276], the structure of the nucleosome core particle in complex with Chd1 was assumed to be unchanged compared to the x-ray crystallographic structure (pdb ID: 3LZ0) [256]. This assumption was tested by distance measurements discussed earlier in this chapter (section 4.4). The distance controls showed that the nucleosome exit site does not alter in the presence of Chd1, however a slight displacement of the DNA at the nucleosome entry site could not be excluded. To account for this possible conformational alteration compared to the crystal structure, satellite R(-66) was given a larger position prior (i.e. 7 Å longer linker, see Table 4.6.4) and nevertheless, the structure of the 601 nucleosome was used as a reference

frame. Accordingly, the satellite attachment points are known from the crystal structure, whereas the precise location and orientation of the dye molecules are not. For NPS analysis, we therefore simulated the accessible volumes of the dye molecules by using a flexible linker model, given the point of attachment, the size of the dye molecules and the length and diameter of the linker [20]. All parameters used for the simulation of satellite dye position priors are given in Table 4.6.4. Antenna dye molecules were attached via flexible linkers to positions on the Chd1 N-terminal tandem chromodomains (C190, C293), on the ATPase/helicase domain (C365, C647, C732) or on the C-terminal DNA-binding domain (DBD) (C1091, C1149, C1159). All antenna attachment points were unknown in the coordinate system of the nucleosome and were to be determined within by application of the NPS. However, for the antennas on the DBD of Chd1, we used information available about the local macromolecular structure of the Chd1 DBD in complex with DNA (pdb ID: 3TED) [150], assuming it is not subject to a conformational change when incorporated into a Chd1-nucleosome complex, and simulated accessible volumes relative to this structure in a similar way like for the satellite dyes (Table 4.6.4). By displacing the structure of the DBD-DNA complex also the antenna dyes on the DBD were moved to other positions so that the structure served as a reference frame to the DBD antenna dyes. The final positions and orientations of the antennas were calculated from their positions and orientations relative to the local structure and from the position and orientation of the reference frame attached roughly to the center of mass of the local structure. The volume accessible to the origin of the docked DBD-DNA reference frame was set to a cube of 300 Å x 300 Å x 350 Å with the origin at (x,y,z) = (-20, -20, 100) excluding the volume of the nucleosome. The orientation of the DBD-DNA reference frame remained completely free. The antennas attached to the tandem chromodomains and the ATPase domain of Chd1 were treated in a different way because the conformation these domains adopt when bound to a nucleosome is expected to deviate from their conformation in the available crystal structure of a chromodomain-ATPase domain construct [147]. Therefore, the position priors of the antennas attached to the chromodomains and the ATPase domain were only restricted to a volume not occupied by the nucleosome. To this end, I assumed zero probability density within the volume occupied by the nucleosome shrunk by 2 Å and equal probability elsewhere in a cubic box around the nucleosome with a side length of 300 Å. In addition, distance constraints were applied for the antennas attached to the ATPase domain since the magnitude of conformational change within this domain could be well estimated based on available structural data [147, 148]. The maximal possible distances between the antenna attachment sites were derived from the crystal structure of the chromodomain-ATPase domain portion of Chd1 [147], in which the ATPase domain captures an open conformation. As a result, the distance between Chd1-C365 and C647 was restricted to a range of 0-88 Å, the distance between Chd1-C365 and -C732 to a range of 0-78 Å and the distance between Chd1-C647 and -C732 to a range of 0-65 Å.

Table 4.6.4.: Prior information of satellites and antennas for global NPS analysis of a Chd1-nucleosome complex. For each satellite the coordinates x,y and z of the dye attachment point are given in Angstrom in the nucleosome coordinate system. Further, for the antenna assigned to the reference frame of the DNA-binding domain of Chd1, the atom number, atom name and residue number of the attachment point is given in the reference frame. Additionally, the parameters of the flexible linker and the dye used for modeling the prior are listed: L_{linker} : length of flexible linker, d_{linker} : diameter of the linker, d_{dye} : diameter of the dye.

Satellites							
Dye position	Ref frame (pdb ID)	attachment point			L_{linker} [Å]	d_{linker} [Å]	d_{dye} [Å]
		x	y	z			
F(+14)	3LZ0	-64.557	-16.537	60.635	13	4.5	7
F(+2)	3LZ0	-30.758	-28.976	81.970	13	4.5	7
F(-15)	3LZ0	22.869	-32.583	64.336	13	4.5	7
F(-64)	3LZ0	-59.400	-36.615	57.344	13	4.5	7
R(-60)	3LZ0	24.691	-3.526	55.441	13	4.5	7
R(-66)	3LZ0	12.446	-12.753	70.623	20	4.5	7

Antennas							
Dye position	Ref frame (pdb ID)	attachment point			L_{linker} [Å]	d_{linker} [Å]	d_{dye} [Å]
		chain	residue #	atom			
Chd1-C1091	3TED moved to (0,0,0)	A	1091	OG	7	4.5	7
Chd1-C1149	3TED moved to (0,0,0)	A	1149	OG	7	4.5	7
Chd1-C1159	3TED moved to (0,0,0)	A	1159	OG	7	4.5	7

Global determination of positions and orientations of Chd1 antennas and reference frames by NPS analysis

For details on global NPS analysis, see section 2.5.3. Briefly, Bayesian parameter estimation was applied using all measured average FRET efficiencies with an error of 5 % for all measurements (larger error compared to NPS analysis of Pol II OC, due to relatively low statistics of measurements), as well as all measured steady-state fluorescence anisotropies and isotropic Förster distances R_0^{iso} , to simultaneously infer position and orientation of all fluorophores within the "laboratory" coordinate system of the nucleosome core particle containing the 601 positioning sequence. Moreover, the information of the antennas C1091, C1149 and C1159 linked to the Chd1-DBD reference frame could be used to also position and orient the reference frame itself with respect to the nucleosome coordinate system. As a result, we obtain as posterior the three dimensional probability density function (PDF) of the positions and orientations of antenna and satellite dyes. As a feature of the docking NPS analysis, the marginal position PDF of any desired point (e.g. antenna attachment point) in the docked reference frames relative to the nucleosome can be extracted in addition to the antenna dye position PDFs. The complete calculation was performed using a custom implementation of the nested sampling algorithm [224] based on Markov chain Monte-Carlo in C and MATLAB (*The MathWorks*). The NPS software is available free of charge at www.uni-ulm.de/nawi/nawi-biophys.html [22]. After NPS analysis, the marginal position PDFs of the antenna dyes and of the attachment

points of the dyes within the docked reference frames relative to the nucleosome core particle coordinate system were exported as `XPLOR`- or `mrc`-files with a resolution of 8-12 Å. The credible volumes, i.e. the three-dimensional error bars, of the positions were finally displayed as isosurfaces in Chimera [151].

Determination of isotropic Förster distances and anisotropies

Isotropic Förster distances and steady state fluorescence anisotropies were determined following the procedure described in section 3.6.8. For quantum yield determination of the donor dye Tamra attached to nucleosomal DNA, DNA was prepared by PCR (see section 4.6.4) and purified by using a PCR purification kit (*Qiagen*). The resulting semi-crude DNA was directly used for acquisition of absorption and fluorescence emission spectra for quantum yield determination. Steady-state fluorescence anisotropies were measured using either Tamra-labeled nucleosomes diluted to a final concentration of 10-60 ng/ μ l (referring to DNA mass) in a total volume of 65 μ l or using Alexa647-labeled Chd1 mutants (about 2 pmol) bound to unlabeled nucleosomes (4-10 pmol) in the presence of 150 μ M AMP-PNP in a total volume of 65 μ l.

5. Outlook

This thesis focused on two fundamental processes of eukaryotic gene expression: transcription initiation carried out by the DNA-dependent RNA polymerase II (Pol II) and nucleosome repositioning by the ATP-hydrolyzing chromatin remodeling enzyme Chd1. Both processes are tightly connected *in vivo*, because RNA polymerase and associated general transcription factors need to access genomic DNA in order to initiate transcription and this accessibility, which is hindered due to the compaction of DNA into nucleosomes, is regulated by chromatin remodelers. The overall aim was to elucidate the molecular mechanisms underlying both processes and to solve the structural architecture of the responsible multi-subunit macromolecular complexes. Since transcription initiation and nucleosome remodeling, are characterized by large conformational changes and a high flexibility of the involved nucleic acid-protein complexes, analysis by means of standard high-resolution ensemble-based structural methods, e.g. x-ray crystallography and NMR spectroscopy, are hindered and instead direct visualization in real time is required as provided by single molecule techniques such as single molecule FRET.

In the first part, I used smFRET experiments, NPS analysis and modeling based on x-ray crystallographic information to determine the three-dimensional architecture of a minimal Pol II open promoter complex consisting of promoter DNA including a TATA box and a mismatched bubble around the transcription start site, TBP, Pol II and general transcription factors TFIIB and TFIIF. The results uncovered large overall structural changes during the initiation-elongation transition, which are apparently accommodated by the intrinsic flexibility of TFIIB. Moreover, I directly observed the downstream DNA to be dynamically loaded into and unloaded from the Pol II cleft at a timescale of seconds. TFIIF was observed to influence these dynamics by stabilizing the downstream DNA in the cleft.

In the second part, I applied the same experimental approach to determine the location of the three constituting domains of chromatin remodeler Chd1 relative to a nucleosome in a Chd1-nucleosome complex that showed intrinsic dynamics. The results of NPS analysis allowed me to construct a model of the Chd1-nucleosome complex that provided insights into the mechanism of Chd1 remodeling. Furthermore, I used smFRET to follow the structural dynamics of nucleosomal DNA during Chd1-catalyzed repositioning in real time. The FRET time trajectories revealed gradual and bidirectional translocation of nucleosomal DNA by Chd1 on a time scale of hundreds of milliseconds to seconds, and the data allowed me to propose a model for the remodeling mechanism of Chd1, which

involves formation and propagation of a DNA loop.

In summary, the results of this thesis present a substantial progress in our understanding of the molecular mechanisms of transcription initiation and nucleosome repositioning and they demonstrate that smFRET combined with NPS analysis is an extremely powerful tool to study the structure as well as the dynamics of large, flexible and even heterogeneous nucleic acid-protein complexes not amenable to other structural and biochemical ensemble techniques. Considerations about required follow-up experiments and possible future directions of both research projects individually were already provided in the "summary and outlook" section of the respective chapters (sections 4.5 and 3.5). Here I want to give a more general outlook.

Despite the progress described here, all experiments constituting this work hold one major drawback: they were performed *in vitro* and the studied enzymatic machines were placed into an extremely controlled environment not resembling their natural place of action. Hence, it remains to be seen how the simplistic mechanistic picture obtained from my *in vitro* studies has to be adapted to describe the complexity within organisms. Along this line, I propose that future single molecule research of both, transcription and chromatin remodeling, should be directed towards the exploration of more complex and hence more realistic situations.

More concretely, future single molecule research of transcription initiation should move away from the study of minimal transcription initiation intermediates containing artificial promoters and instead investigate initiation on natural promoters in the presence of all naturally occurring general transcription factors, including TFIIA, the full TFIID complex including TBP and TAFs, TFIIE and TFIIH in addition to the here used factors TBP, TFIIB and TFIIF. In the presence of the helicase TFIIH, one could directly observe opening of the promoter in real time during the closed to open complex transition and test if the mechanistic picture gained in this thesis still applies. Moreover, Pol II does not encounter bare DNA *in vivo*, but a large portion of DNA is occluded through compaction in nucleosomes. Therefore, transcription initiation and elongation should be examined on a DNA template containing nucleosomes in order to learn how Pol II reacts on and transcribes through these obstacles. In order to study more complex and more natural transcription initiation scenarios by smFRET and NPS analysis, it will be absolutely essential to site-specifically label the involved transcription factors as well as Pol II itself, which has not been possible until now using traditional cysteine chemistry. Due to the fast development of new labeling approaches that allow efficient and site-specific labeling *in vivo* and *in vitro* based on the introduction of genetically encoded peptide tags or unnatural amino acids into the sequence of the desired protein (see section 2.1.5, [188, 190, 191, 186, 187]), it will soon be possible to site-specifically introduce fluorescent labels even into Pol II. This will pave the way to the exploration of architecture and

function of more complex and at the same time more realistic transcription initiation intermediates.

To apply NPS analysis to such larger complex assemblies, it will be helpful to extend the hybrid character of this approach by combining it with other low resolution methods for structure determination such as small-angle x-ray scattering (SAXS), electron microscopy (EM), pulsed dipolar electron spin resonance (ESR) or biochemical crosslinking experiments. SAXS and EM can provide information about the overall shape of a macromolecular complex, however it is difficult to assign different parts of a density to the individual domains of the examined complex. Hence, NPS is complementary to these techniques and the combination of methods would greatly increase the accuracy of resulting structural models. Moreover, ESR can deliver, similarly to smFRET, long range distance information in solution, whereas biochemical crosslinking provides information about residues in close proximity. Both methods can therefore add information to the global NPS analysis.

Concerning nucleosome remodeling, single molecule studies should not limit themselves to exploring the repositioning of mononucleosomes assembled on the strong nucleosome positioning sequence 601. Remodelers never encounter single nucleosomes in the natural environment of a living cell and especially for the spacing activity of Chd1 or ISWI remodelers, the presence of neighboring nucleosomes might be essential and have a strong effect on the applied remodeling mechanism. The importance of a dinucleosome substrate has already been shown for ISW1a in a recent x-ray and electron microscopy structural study [154]. Moreover, nucleosomes containing the 601 sequence pose an artificial substrate to remodelers due to the strong periodicity of stable DNA-histone interactions and due to the overall unnaturally high stability. Hence, future studies should explore remodeling of nucleosomes assembled on different DNA substrates and, furthermore, should use nucleosomal arrays or stretches of chromatin fiber as substrates. For the analysis of di-nucleosome remodeling, application of three-color FRET might be adequate, however for the examination of remodeling of higher order chromatin structures, the application of smFRET, even with three colors, is not feasible anymore. Here, the fast emerging super-resolution fluorescence imaging techniques, such as STORM [277], based on the localization of single molecules, or STED [278], based on spatially patterned excitation, will be the method of choice, at best in a multi-color mode. If nucleosomal arrays or chromatin fibers are labeled at the histones and with a different color at the DNA, structural dynamics and nucleosome repositioning as a result of the action of a chromatin remodeler can be visualized using these super-resolution microscopic techniques.

Since transcription and chromatin remodeling are tightly connected *in vivo*, and chromatin remodelers regulate transcription by controlling accessibility to the genomic DNA, one further future goal should be to combine chromatin remodeling and transcription initiation by following both processes consecutively on a single substrate. Transcription of

a nucleosomal template by individual Pol II enzymes should be examined in the presence and absence of chromatin remodeling enzymes such as Chd1, which has been affiliated with transcription initiation and elongation *in vivo* [136, 132, 133, 134], in order to learn about the interplay between both enzymes. Such experiments will provide a starting point for the investigation of more complex regulatory networks.

Finally, the application of single molecule fluorescence techniques is not limited to *in-vitro* studies. Fast advances in live-cell imaging technologies have made it possible to apply single molecule fluorescence microscopy to cells and directly observe individual transcription events in single living cells (see section 1.1.4, [279, 280]). Hence, transcription initiation, chromatin remodeling and the interplay between both processes should ultimately be investigated in the natural environment of a living cell or even of a living organism such as zebrafish, suited for fluorescence microscopic investigation due to its transparency during early development. Nascent RNA transcripts can be labeled fluorescently in living cells by using the MS2-stem loop technique [93] and labeling of proteins can be achieved in multiple ways, due to the fast development of new labeling techniques. One possibility is to create a fusion construct of a fluorescent protein with the protein of interest, e.g. Pol II, that is then expressed in the fluorescently tagged version in the investigated living cell. Another possibility is the introduction of a peptide tag or an unnatural amino acid genetically into the protein, which can post-translationally be modified by a fluorescent label in a bio-compatible chemical reaction [188, 190, 191]. A challenge of imaging single molecules in live cells is the strong cellular autofluorescence. TIRF microscopy can help minimizing this autofluorescence, but due to the limited penetration depth of the evanescent field it is not suited to image single molecules in eukaryotic cells, which have a larger cell body compared to bacteria. Hence, new methods such as light sheet microscopy might have to be applied in order to be able to visualize single transcription or remodeling events in living eukaryotic cells or organisms [281, 282, 283].

To conclude, I am confident of having convinced the reader with this thesis of the great power of single molecule fluorescence techniques in answering open questions in biology, especially on the mechanistic and structural level. I am looking forward to reading about many more exciting discoveries that single molecule FRET and the Nano-Positioning System will entail in the future.

Appendix A. Supplemental Data

A.1. Supplemental data for global NPS analysis of Pol II OC

A Supplemental Data

Antenna	Satellite	Anisotropy (Donor)	Anisotropy (Acceptor)	QY [%]	$R_{0,iso}$ [Å]	# of molecules	FRET efficiency	% of probability
Localization of NT-DNA(-18)								
NT-DNA(-18)-T	Rpb7-C150-A	0.26	0.25	25	59	128	49	-
NT-DNA(-18)-T	T-DNA(-10)-A	0.26	0.3	25	58	186	56*/83	19/81
NT-DNA(-18)-T	T-DNA(+3)-A	0.26	0.32	25	58	480	56/81*	67/33
NT-DNA(-18)-T	T-DNA(+7)-A	0.26	0.33	25	59	192	37/83*	62/38
NT-DNA(-18)-T	T-DNA(+12)-A	0.26	0.34	25	57	172	42/79*	64/36
Localization of NT-DNA(-23)								
NT-DNA(-23)-T	Rpb7-C150-A	0.28	0.25	36	65	49	33	-
NT-DNA(-23)-T	T-DNA(-10)-A	0.28	0.3	36	60	447	66	-
NT-DNA(-23)-T	T-DNA(+3)-A	0.28	0.32	36	60	382	31/64*	84/16
NT-DNA(-23)-T	T-DNA(+7)-A	0.28	0.33	36	61	346	33/72*	64/36
NT-DNA(-23)-T	T-DNA(+12)-A	0.28	0.34	36	61	576	33/74*	81/19
Localization of NT-DNA(-30)								
NT-DNA(-30)-T	Rpb7-C150-A	0.3	0.25	26	62	70	12	-
NT-DNA(-30)-T	T-DNA(-10)-A	0.3	0.3	26	57	377	48/73*	79/21
NT-DNA(-30)-T	T-DNA(+3)-A	0.3	0.32	26	57	456	29/75*	61/39
NT-DNA(-30)-T	T-DNA(+7)-A	0.3	0.33	26	58	349	43/80*	66/34
NT-DNA(-30)-T	T-DNA(+12)-A	0.3	0.34	26	56	587	16/81*	89/11
Localization of TBP-C128								
TBP-C128-A	Rpb7-C150-C	0.30	0.25	20	57	139	32/53*/81*	53/36/11
TBP-C128-A	T-DNA(-10)-T	0.32	0.25	31	60	199	38/70*	71/29
TBP-C128-A	T-DNA(+3)-T	0.30	0.25	36	62	159	32	-
TBP-C128-A	T-DNA(+7)-T	0.31	0.25	33	61	129	19	-
TBP-C128-A	T-DNA(+12)-T	0.22	0.25	28	61	110	17	-
TBP-C128-A	NT-DNA(-23)-T	0.28	0.25	36	62	91	45*/87	44/56
TBP-C128-A	NT-DNA(-30)-T	0.30	0.25	26	59	179	50*/85	43/57
Localization of TFIIIB-(122)ybbR								
TFIIIB-(122)ybbR-A	Rpb7-C150-C	0.30	0.25	20	55	104	29/51*	61/39
TFIIIB-(122)ybbR-A	T-DNA(-10)-T	0.32	0.25	31	58	127	53	-
TFIIIB-(122)ybbR-A	T-DNA(+3)-T	0.30	0.25	36	60	247	53/79*	86/14
TFIIIB-(122)ybbR-A	T-DNA(+7)-T	0.31	0.25	33	59	85	52	-
TFIIIB-(122)ybbR-A	T-DNA(+12)-T	0.22	0.25	28	59	99	49/21*	67/33
TFIIIB-(122)ybbR-A	NT-DNA(-23)-T	0.28	0.25	36	60	53	55	-
Localization of TFIIIB-(C-term)ybbR								
TFIIIB-(C-term)ybbR-A	Rpb7-C150-C	0.30	0.23	20	55	22	11	-
TFIIIB-(C-term)ybbR-A	T-DNA(-10)-T	0.32	0.23	31	58	45	34/75*	91/09
TFIIIB-(C-term)ybbR-A	T-DNA(+3)-T	0.30	0.23	36	60	24	21/83*	88/12
TFIIIB-(C-term)ybbR-A	T-DNA(+7)-T	0.31	0.23	33	59	24	20	-
TFIIIB-(C-term)ybbR-A	T-DNA(+12)-T	0.22	0.23	28	59	16	11	-
Localization of satellites within OC								
T-DNA(-10)-T	Rpb7-C150-A	0.32	0.25	31	63	34	44	-
T-DNA(+3)-T	Rpb7-C150-A	0.30	0.25	36	65	114	44/ 65*	57/43
T-DNA(+7)-T	Rpb7-C150-A	0.31	0.25	33	64	135	32/ 57*	57/43
T-DNA(+12)-T	Rpb7-C150-A	0.22	0.25	28	61	213	28/ 50*	67/23
Localization of satellites within EC								
T-DNA(-10)-T	Rpb7-C150-A	0.32	0.25	31	63	70	41	-
T-DNA(+3)-T	Rpb7-C150-A	0.30	0.25	36	65	23	45	-
T-DNA(+7)-T	Rpb7-C150-A	0.31	0.25	33	64	176	37	-
Localization of NT-DNA(-23) within OC lacking TFIIIF								
NT-DNA(-23)-T	Rpb7-C150-A	0.28	0.25	36	65	121	33	-
NT-DNA(-23)-T	T-DNA(-10)-A	0.28	0.3	36	60	1083	63	-
NT-DNA(-23)-T	T-DNA(+3)-A	0.28	0.32	36	60	489	36/88	50/50
NT-DNA(-23)-T	T-DNA(+7)-A	0.28	0.33	36	61	300	33/83*	57/43
NT-DNA(-23)-T	T-DNA(+12)-A	0.28	0.34	36	61	797	32*/74	42/58

Table A.1.1.: Overview of experimental data measured for NPS analysis of OCs including control measurements of ECs. (QY: quantum yield, R_0^{iso} : isotropic Förster distance, *: side peak. Abbreviations for fluorophores: A: Alexa 647, T: Tamra, C: Cy3b.)

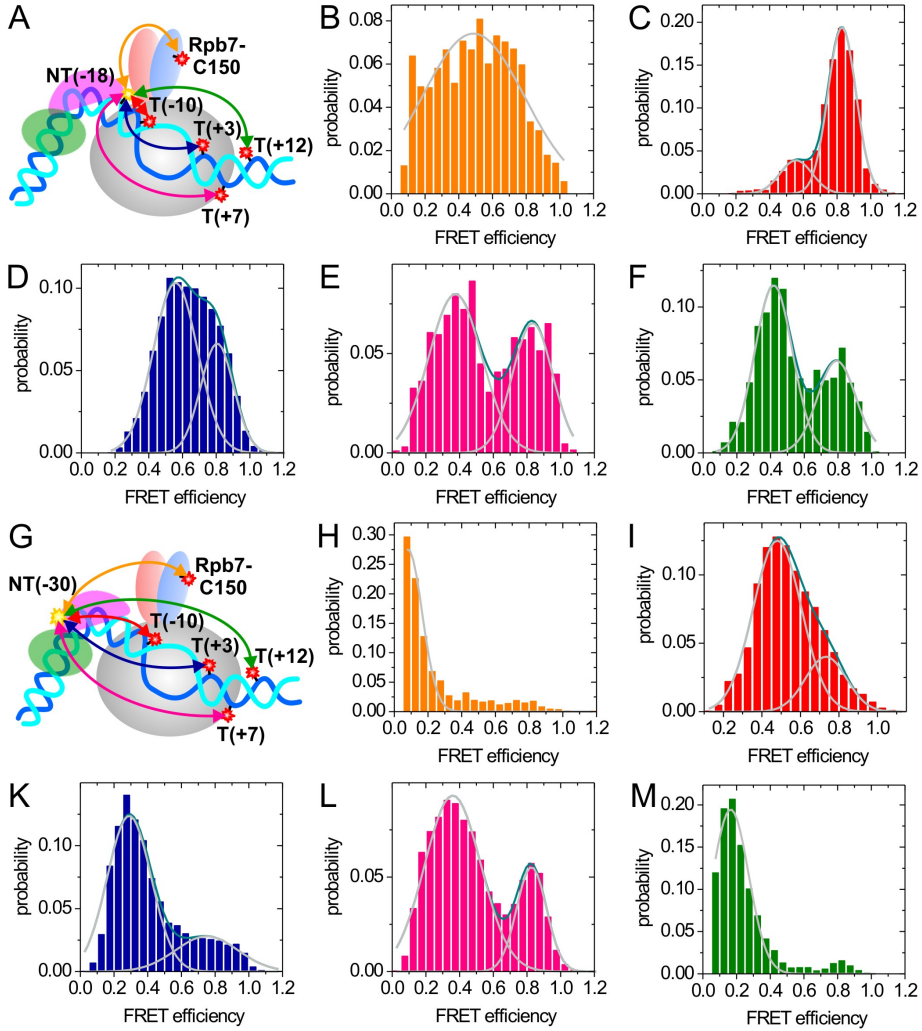


Figure A.1.1.: smFRET measurements for localization of NT-DNA(-18) and NT-DNA(-30). (A) Cartoon illustrating the set of smFRET measurements for the localization of antenna NT-DNA(-18) by NPS. Measurements are indicated by double-headed arrows colored orange for satellite Rpb7-C150, red for T-DNA(-10), blue for T-DNA(+3), pink for T-DNA(+7) and green for T-DNA(+12). An OC is presented schematically using the same colors as in Figure 3.3.1. The same color coding will be used in the following figures. (B-F) smFRET measurements from the five satellites to antenna NT-DNA(-18) on the upstream DNA. Frame-wise FRET efficiency histograms are shown in the color of the respective satellite (see Figure 3.3.1C and color of arrows in panel A): (B) satellite Rpb7-C150, (C) T-DNA(-10), (D) T-DNA(+3), (E) T-DNA(+7), (F) T-DNA(+12). The same color coding will be used in all following histograms. FRET efficiency histograms were fitted with one or two Gaussians (gray, individual Gaussian fit; dark cyan, combined fit). From the fits, the mean FRET efficiencies and their standard errors were determined and are summarized in Table A.1.1. (G) Cartoon illustrating the set of smFRET measurements for the localization of antenna NT-DNA(-30) by NPS. (H-M) Frame-wise FRET efficiency histograms of the measurements from the five satellites to antenna NT-DNA(-30) on the TATA box: (H) satellite Rpb7-C150, (I) T-DNA(-10), (K) T-DNA(+3), (L) T-DNA(+7), (M) T-DNA(+12).

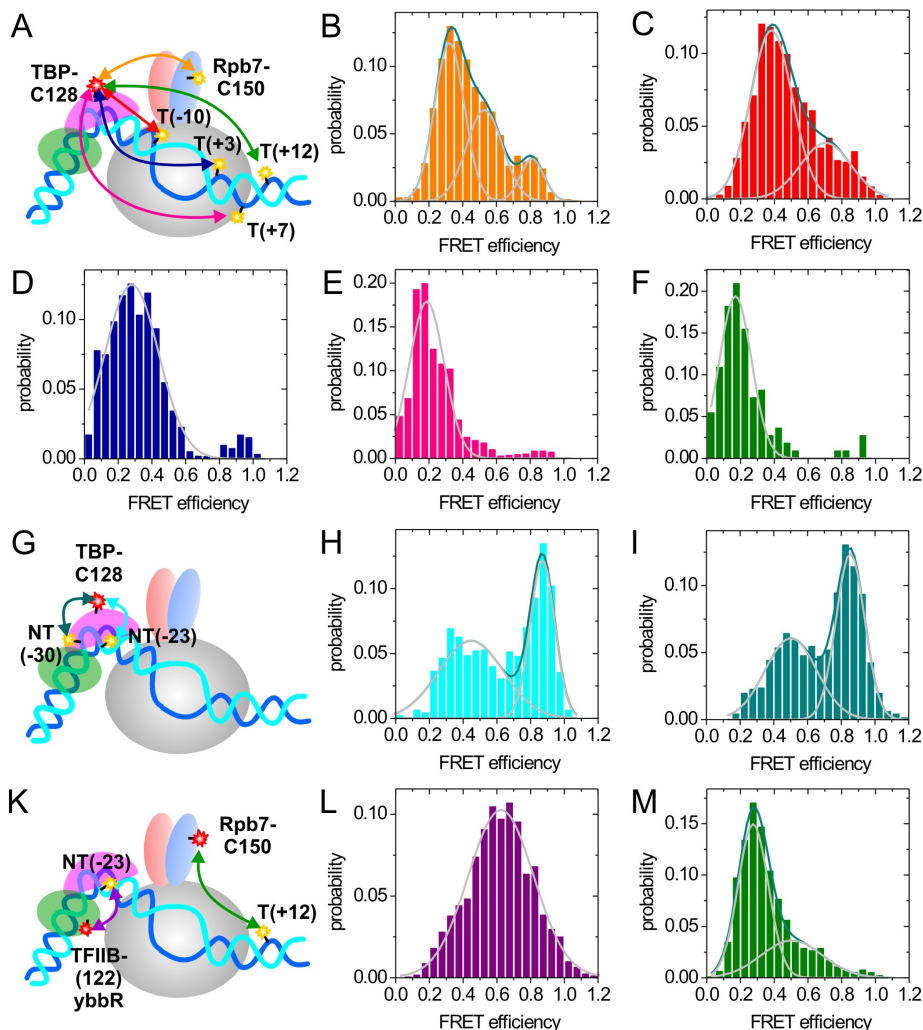


Figure A.1.2.: smFRET measurements for localization of TBP-C128 and distance tests between TBP-C128/TATA, TFIIB-(122)ybbR/TATA and T-DNA(+12)/Rpb7-C150. (A) Cartoon illustrating the set of smFRET measurements for the localization of antenna TBP-C128 by NPS. Measurements are indicated by double-headed arrows in the color of the respective satellite (see Figure 3.3.1). (B-F) Framework FRET efficiency histograms of the measurements from the five satellites to antenna TBP-C128: (B) satellite Rpb7-C150, (C) T-DNA(-10), (D) T-DNA(+3), (E) T-DNA(+7), (F) T-DNA(+12). (G) Cartoon illustrating the smFRET measurements between antenna TBP-C128 and the antennas on the TATA box. (H,I) Framework FRET efficiency histograms of the measurements illustrated in panel G: (H) between TBP-C128 and NT-DNA(-23) and (I) between TBP-C128 and NT-DNA(-30). (K) Cartoon illustrating the smFRET measurement between antenna TFIIB-(122)ybbR and antenna NT-DNA(-23) on the TATA box and between the satellites T-DNA(+12) and Rpb7-C150. (L) Framework FRET efficiency histogram of the measurement between the antennas TFIIB-(122)ybbR and NT-DNA(-23) illustrated in panel K. (M) Framework FRET efficiency histograms of the measurement between the satellites T-DNA(+12) and Rpb7-C150 illustrated in panel K.

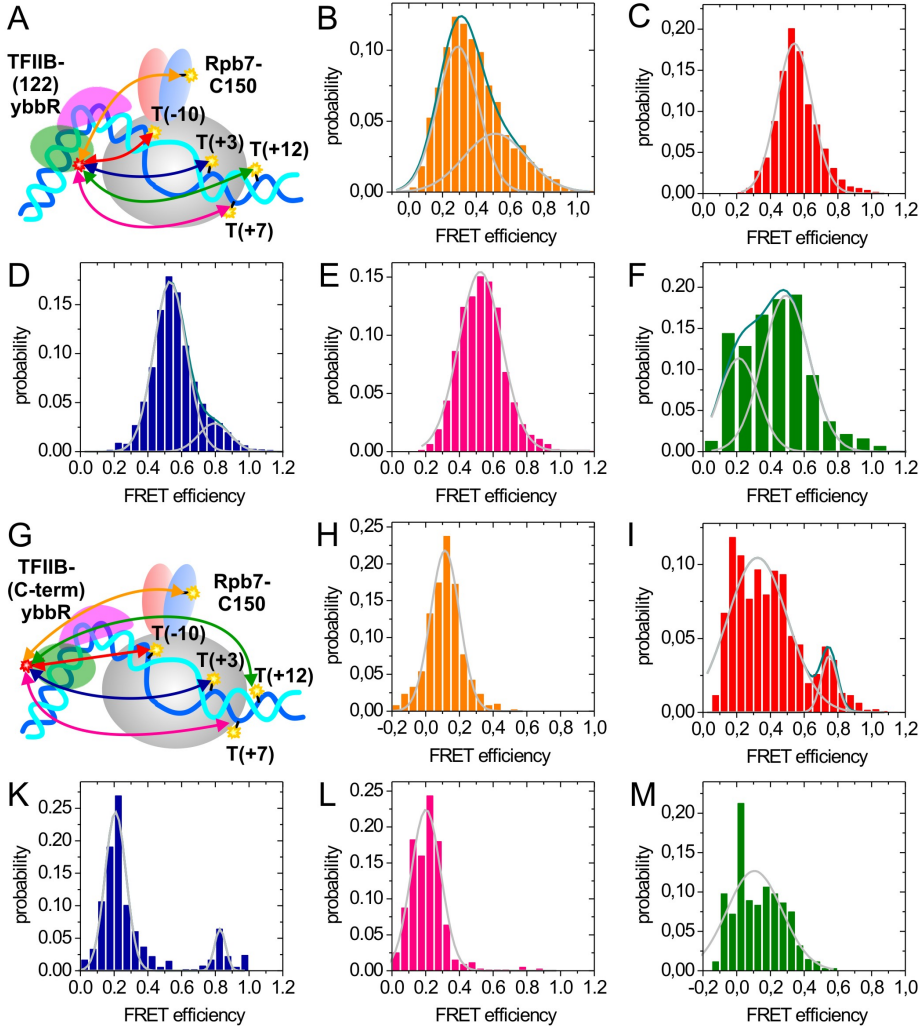


Figure A.1.3.: smFRET measurements for localization of TFIIB-(C-term)ybbR and TFIIB-(122)ybbR. (A) Cartoon illustrating the set of smFRET measurements for the localization of antenna TFIIB-(122)ybbR by NPS. Measurements are indicated by double-headed arrows in the color of the respective satellite (see Figure 3.3.1). (B-F) Framework FRET efficiency histograms of the measurements from the five satellites to antenna TFIIB-(122)ybbR: (B) satellite Rpb7-C150, (C) T-DNA(-10), (D) T-DNA(+3), (E) T-DNA(+7), (F) T-DNA(+12). (G) Cartoon illustrating the set of smFRET measurements for the localization of antenna TFIIB-(C-term)ybbR by NPS. Measurements are indicated by double-headed arrows in the color of the respective satellite (see Figure 3.3.1C). (H-M) Framework FRET efficiency histograms of the measurements from the five satellites to antenna TFIIB-(C-term)ybbR: (H) satellite Rpb7-C150, (I) T-DNA(-10), (K) T-DNA(+3), (L) T-DNA(+7), (M) T-DNA(+12).

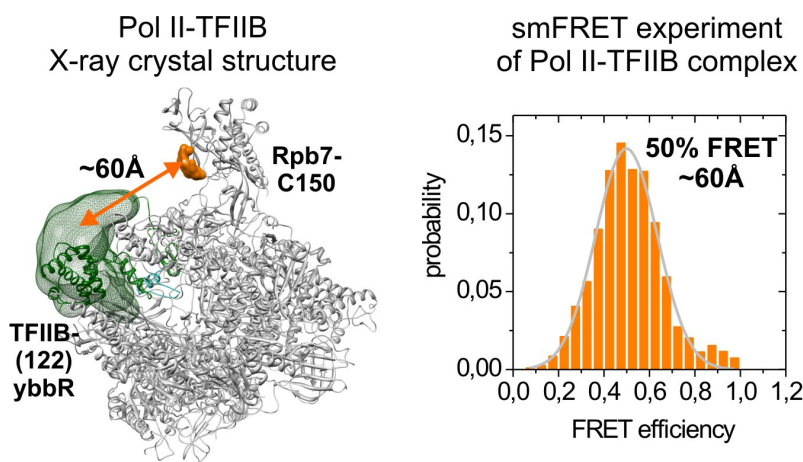


Figure A.1.4.: Pol II-TFIIB complex captures a similar structure in smFRET measurement conditions and in x-ray crystallography. smFRET was measured between Rpb7-C150 and TFIIB-(122)ybbR in Pol II-TFIIB complexes and the distance resulting from the FRET efficiency histogram with a mean FRET efficiency of 50 % (right) was compared to the distance in the x-ray structure (left). Both methods revealed a similar distance of about 60 Å (assuming the isotropic R_0 value) showing that Pol II-TFIIB complexes capture the same structure in smFRET measurement conditions as in the crystal structure. Note that the corresponding smFRET histogram of the open complex (Figure A.1.3B) looks dramatically different and shows a main peak with a FRET efficiency of about 30 %.

A.2. HMM analysis of dynamic FRET traces during Chd1 remodeling

For each set of dynamic smFRET trajectories of Chd1 remodeling of a given nucleosome sample, different HMM models with different parameter settings were tested to find the model that best describes the data. Therefore, HMM analysis was performed with different number of states N either globally with one parameter set to describe all trajectories of a given sample or locally with an individual parameter set for each trajectory. The number of states was estimated either by eye or with the help of the software *vbFRET* (variational bayesian inference for smFRET time series) [275], which uses Bayesian inference to identify the most probable number of states present in each dynamic time trace and thus provides a histogram of number of states of a given data set. Moreover, the width σ of states was either kept fixed at one value (usually values between 0.01 and 0.2 were tested) or was allowed to be learned during model training. The following two figures illustrate the model testing procedure for the dynamic smFRET time traces recorded during remodeling of R(-85)/F(+14) nucleosomes in the presence of 50 nM Chd1 and 150 μ M ATP. Transition density plots (TDPs) for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from HMM analyses are shown.

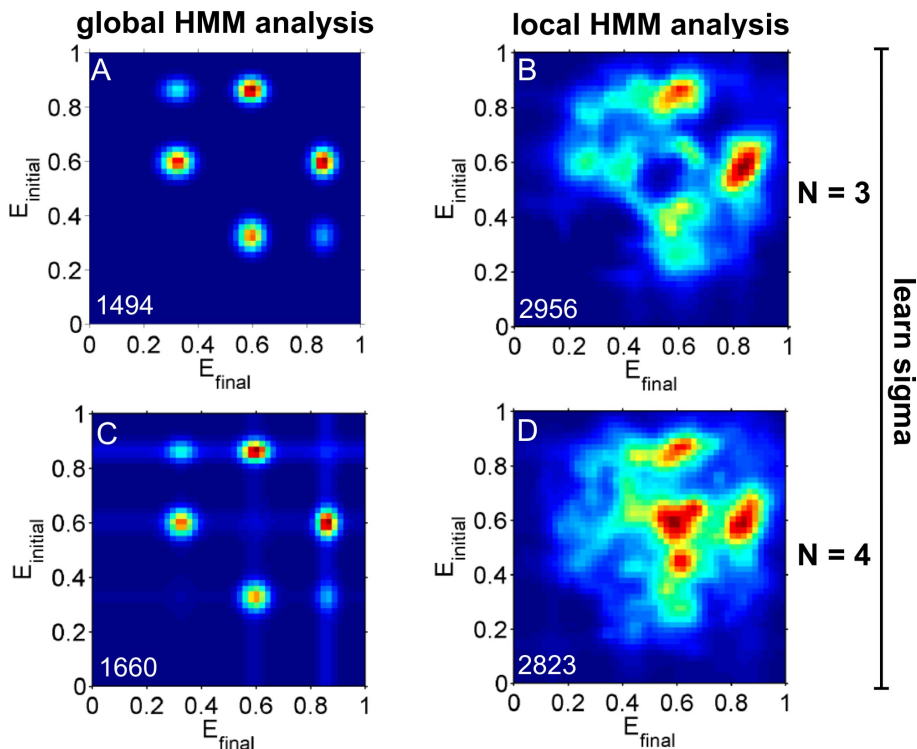


Figure A.2.1.: Test of different Hidden Markov Models to describe the dynamic smFRET trajectories (378 trajectories) observed during remodeling of R(-85)-Alexa 647/ F(+14)-Tamra nucleosomes in the presence of 50 nM Chd1 and 150 μ M ATP. TDPs for FRET efficiency transitions from the initial FRET efficiency E_{initial} to the final FRET efficiency E_{final} resulting from the different HMM analyses are shown with the number of detected transitions in the lower left corner of each TDP in white. HMM analysis was performed with variable width of states σ (learn σ during model training) either globally (left, A and C) with one parameter set to describe all trajectories or locally (right, B and D) with an individual parameter set for each trajectory (see section 2.4 for details). A three-state HMM (A,B) as well as a four-state HMM (C,D) was tested. Global analysis did not describe well the data since the initial FRET state of R(-85)/F(+14) nucleosomes at $\sim 45\%$ FRET was not detected and even the four-state model found only three distinguishable states. In addition, cumulative dwell time histograms extracted from the TDPs could not well be fitted with mono-exponential decay functions indicating that different pathways were combined into one transition population. Local analysis was able to detect the initial FRET state at $\sim 45\%$, however transitions did not cluster well and were found also on the diagonal (transitions between states with similar FRET efficiency value). Therefore, HMM analysis was tested with a fixed width of FRET states (see Figure A.2.2).

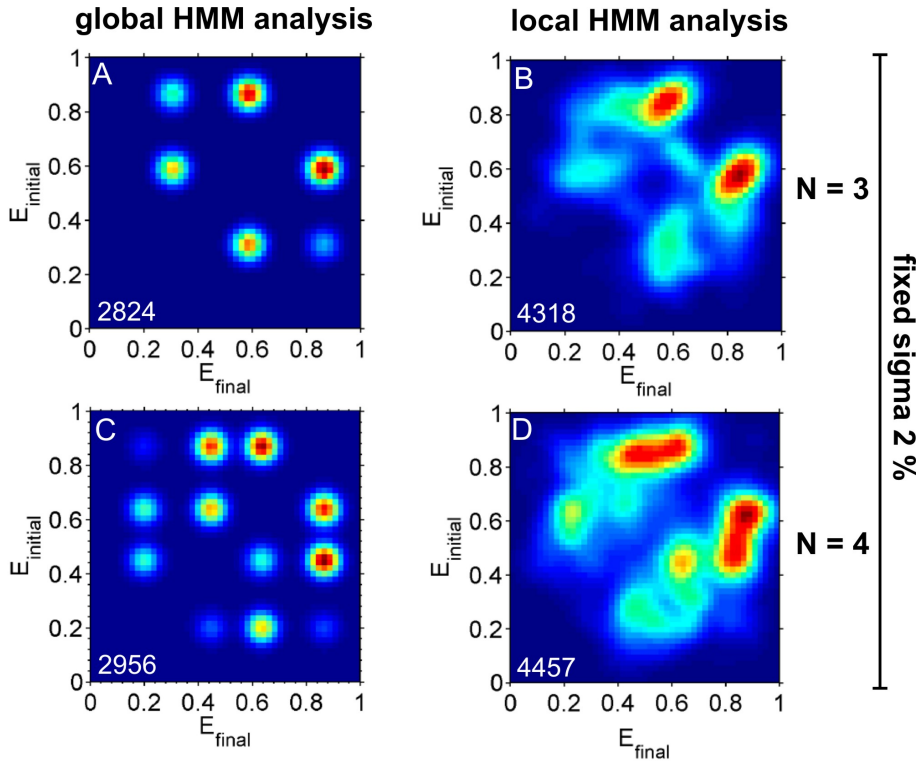


Figure A.2.2.: Test of different Hidden Markov Models to describe the dynamic smFRET trajectories (378 trajectories) observed during remodeling of R(-85)-Alexa 647/ F(+14)-Tamra nucleosomes in the presence of 50 nM Chd1 and 150 μ M ATP. TDPs for FRET efficiency transitions from the initial FRET efficiency $E_{initial}$ to the final FRET efficiency E_{final} resulting from the different HMM analyses are shown with the number of detected transitions in the lower left corner of each TDP in white. HMM analysis was performed with a fixed width of states σ of 0.02 either globally (left, A and C) with one parameter set to describe all trajectories or locally (right, B and D) with an individual parameter set for each trajectory (see section 2.4 for details). A three-state HMM (A,B) as well as a four-state HMM (C,D) was tested. Global three-state HMM analysis (A) did not describe well the data, since the initial FRET state of R(-85)/F(+14) nucleosomes at ~ 45 % FRET was not detected. In addition, cumulative dwell time histograms extracted from the TDPs could not well be fitted with mono-exponential decay functions indicating that different pathways were combined into one transition population. However, global four-state HMM analysis (C) yielded a good result and detected all FRET states (including the initial FRET state) that were also visible by eye in the dynamic FRET time traces. Moreover, cumulative dwell time histograms could well be described by mono-exponential decays (see Figure A.2.3 for complete set of cumulative dwell time histograms). Therefore, this four-state HMM analysis with $\sigma = 0.02$ was used for further interpretation of the data (see section 4.3.1). Local four-state HMM analysis yielded reasonable results as well, however transition populations were less well resolved. Since global analysis is more conform with the idea of HMM analysis and model training, global analysis was preferentially applied for analysis of dynamic smFRET data of nucleosome remodeling.

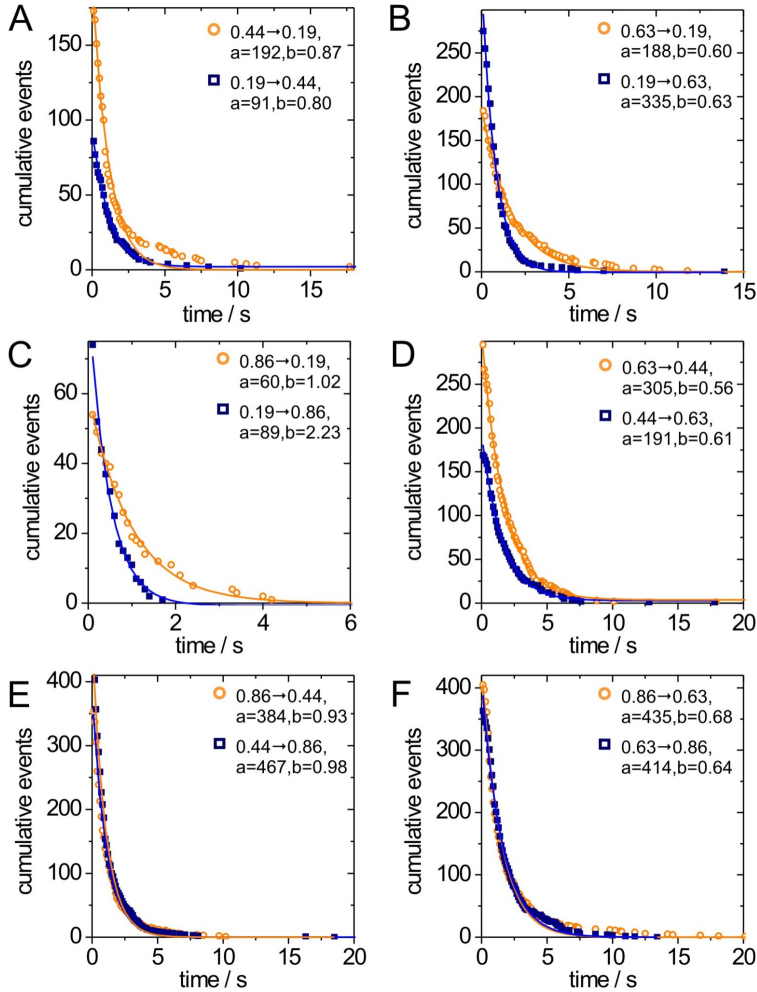


Figure A.2.3.: Cumulative dwell time distributions showing the number of transition events with a dwell time longer than a given time for the transitions observed during remodeling of R(-85)-Alexa 647/ F(+14)-Tamra nucleosomes in the presence of 50 nM Chd1 and 150 μ M ATP (see Figure 4.3.4). The forward and backward transitions between two distinct states are shown in one figure with the transition from high to low FRET as orange circles and the transition from low to high FRET as blue squares. The solid lines represent the corresponding mono-exponential decay fits $y = ae^{-bx}$ and the fit parameters a and b are given in each legend. (A) Transitions between 44 % and 19 % FRET, (B) 63 % and 19 % FRET, (C) 86 % and 19 % FRET, (D) 63 % and 44 % FRET, (E) 86 % and 44 % FRET, (F) 86 % and 63 % FRET.

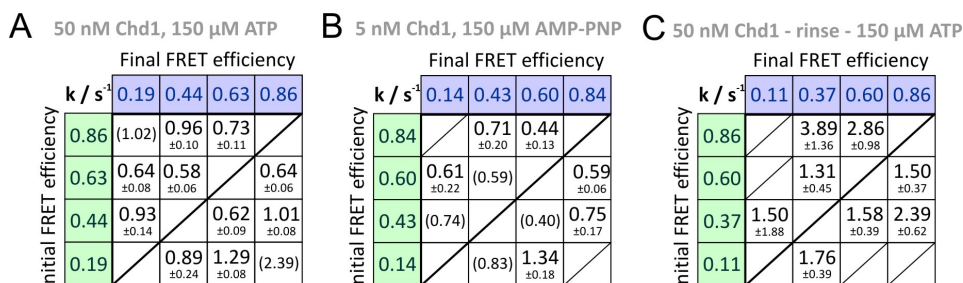


Figure A.2.4.: Kinetic rates of transitions between individual FRET states observed during Chd1-catalyzed remodeling of R(-85)-Alexa 647/ F(+14)-Tamra nucleosomes as obtained by global HMM analysis of dynamic FRET traces. Each matrix presents transition rates and their standard deviations in s^{-1} for all detected transitions for a given remodeling condition: (A) 50 nM Chd1, 150 μ M ATP; (B) 5 nM Chd1, 150 μ M ATP; (C) 50 nM Chd1, removal of excess Chd1 and subsequent addition of 150 μ M ATP. Each rate was extracted from a mono-exponential fit to the cumulative distribution of dwell times of each transition. The error is given as standard deviation of the rate from randomly chosen subsets of 30 % of molecules undergoing the same transition. Rates in parenthesis have only low statistics due to a low occurrence of the described transition.

A.3. Supplemental data for global NPS analysis of Chd1-nucleosome complex

Antenna	Satellite	Anisotropy (Acceptor)	Anisotropy (Donor)	QY [%]	$R_{0,iso}$ [Å]	# of molecules	FRET efficiency	% of probability
Localization of Chd1 tandem chromodomains								
Chd1-C190-A	F(-15)-T	0.11	0.19	36	65	89	31*/87	27/73
Chd1-C190-A	F(+2)-T	0.11	0.27	31	64	87	40/51*	75/25
Chd1-C190-A	R(-60)-T	0.11	0.24	42	67	118	23/35*	72/28
Chd1-C190-A	R(-66)-T	0.11	0.28	46	68	101	38*/51	28/72
Chd1-C293-A	F(-15)-T	0.12	0.19	36	65	78	21/54*/84*	70/23/7
Chd1-C293-A	F(+2)-T	0.12	0.27	31	64	40	29*/48	32/68
Chd1-C293-A	F(+14)-T	0.12	0.21	36	65	43	18/62*	79/21
Chd1-C293-A	R(-60)-T	0.12	0.24	42	67	33	15/30*	74/26
Chd1-C293-A	R(-66)-T	0.12	0.28	46	68	44	32/72*	68/32
Localization of Chd1 ATPase domain								
Chd1-C365-A	F(-64)-T	0.16	0.29	34	64	136	32*/74	26/74
Chd1-C365-A	F(-15)-T	0.16	0.19	36	65	67	28*/52	26/74
Chd1-C365-A	F(+2)-T	0.16	0.27	31	64	130	31/59*	72/28
Chd1-C365-A	F(+14)-T	0.16	0.21	36	65	125	20*/46	28/72
Chd1-C365-A	R(-60)-T	0.16	0.24	42	67	134	42*/73	22/78
Chd1-C365-A	R(-66)-T	0.16	0.28	46	68	155	38*/60	23/77
Chd1-C647-A	R(-60)-T	0.12	0.24	42	67	72	31/59*	84/16
Chd1-C647-A	R(-66)-T	0.12	0.28	46	68	65	33/50*	76/24
Chd1-C732-A	F(-64)-T	0.18	0.29	34	64	58	27/100*	93/07
Chd1-C732-A	F(+2)-T	0.16	0.27	31	64	109	31*/55	21/79
Chd1-C732-A	R(-60)-T	0.16	0.24	42	67	139	34	-
Chd1-C732-A	R(-66)-T	0.16	0.28	46	68	119	40/60*	82/18
Localization of Chd1 DNA-binding domain								
Chd1-C1091-A	F(-15)-T	0.21	0.19	36	65	104	32/74*	77/23
Chd1-C1091-A	F(+2)-T	0.21	0.27	31	64	63	30/77*	80/20
Chd1-C1091-A	F(+14)-T	0.21	0.21	36	65	148	28/55*	67/33
Chd1-C1091-A	R(-60)-T	0.21	0.24	42	67	97	21/55*	70/30
Chd1-C1091-A	R(-66)-T	0.21	0.28	46	68	282	32/50*	77/23
Chd1-C1149-A	F(-64)-T	0.18	0.29	34	64	82	21/60*	72/28
Chd1-C1149-A	F(-15)-T	0.18	0.19	36	65	176	34/55*	81/19
Chd1-C1149-A	F(+2)-T	0.18	0.27	31	64	104	29/69*	77/23
Chd1-C1149-A	F(+14)-T	0.18	0.21	36	65	159	21/45*	65/35
Chd1-C1149-A	R(-60)-T	0.18	0.24	42	67	272	26*/52	19/81
Chd1-C1149-A	R(-66)-T	0.18	0.28	46	68	246	47*/64	32/68
Chd1-C1159-A	F(-64)-T	0.12	0.29	34	64	88	28/73*	79/21
Chd1-C1159-A	F(-15)-T	0.12	0.19	36	65	87	28/58*	80/20
Chd1-C1159-A	F(+2)-T	0.12	0.27	31	64	113	34/60*	66/34
Chd1-C1159-A	F(+14)-T	0.12	0.21	36	65	109	25/51*	80/20
Chd1-C1159-A	R(-60)-T	0.12	0.24	42	67	107	40/74*	81/19
Chd1-C1159-A	R(-66)-T	0.12	0.28	46	68	239	40*/63	24/76

Table A.3.1.: Overview of experimental data measured for NPS analysis of Chd1-nucleosome complexes. (QY: quantum yield, R_0^{iso} : isotropic Förster distance, *: side peak. Abbreviations for fluorophores: A: Alexa 647, T: Tamra.)

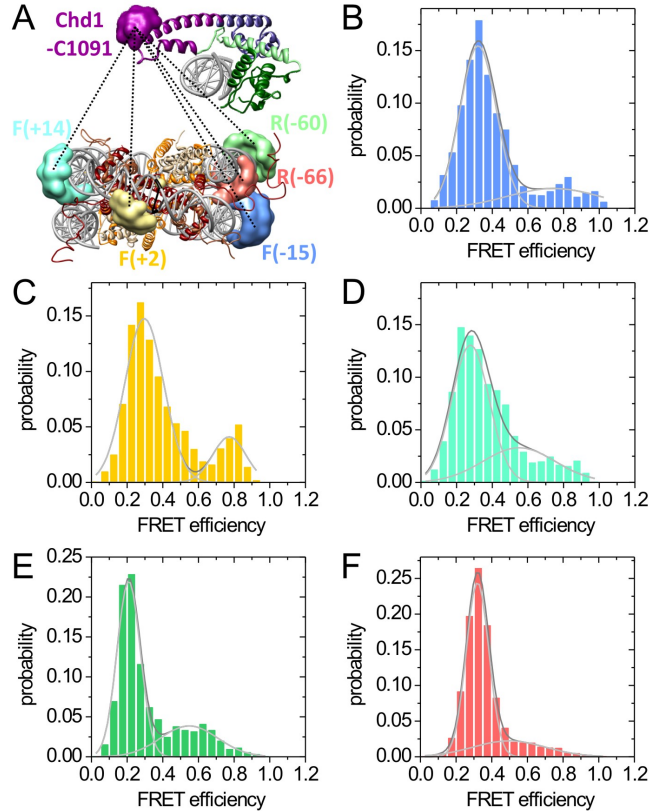


Figure A.3.1.: Positioning of DNA-binding domain of Chd1 relative to the nucleosome core particle. (A) Schematic illustration of the set of smFRET measurements from all satellites to antenna Chd1-C1091 on the DNA-binding domain. The images of nucleosome core particle and Chd1 DNA-binding domain containing the dye position priors were prepared in Chimera [151]. The measurements are indicated by dotted black lines and the respective FRET efficiency histograms are shown in the color of the corresponding satellite (B-F) (light blue for F(-15), yellow for F(+2), aquamarine for F(+14), light green for R(-60), light red for R(-66)). All FRET efficiency histograms were best described with two Gaussians since two populations were visible, one major population comprising 67-80 % of the data and a subpopulation comprising 20-33 % of the data (light gray: individual fits, dark gray: combined fits). The maximum FRET efficiencies of all fits are presented in Table A.3.1.

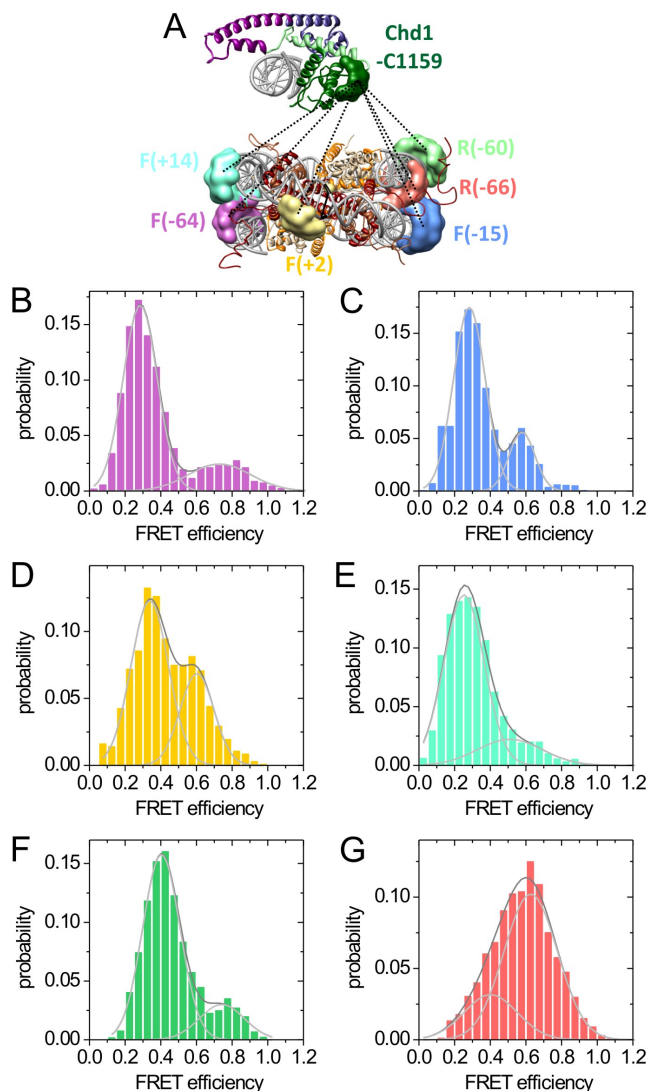


Figure A.3.2.: Positioning of DNA-binding domain of Chd1 relative to the nucleosome core particle. (A) Schematic illustration of the set of smFRET measurements from all satellites to antenna Chd1-C1159 on the DNA-binding domain. The images of nucleosome core particle and Chd1 DNA-binding domain containing the dye position priors were prepared in Chimera [151]. The measurements are indicated by dotted black lines and the respective FRET efficiency histograms are shown in the color of the corresponding satellite (B-G) (magenta for satellite F(-64), light blue for F(-15), yellow for F(+2), aquamarine for F(+14), light green for R(-60), light red for R(-66)). All FRET efficiency histograms were best described with two Gaussians since two populations were visible, one major population comprising 66-81 % of the data and a subpopulation comprising 19-34 % of the data (light gray: individual fits, dark gray: combined fits). The maximum FRET efficiencies of all fits are presented in Table A.3.1.

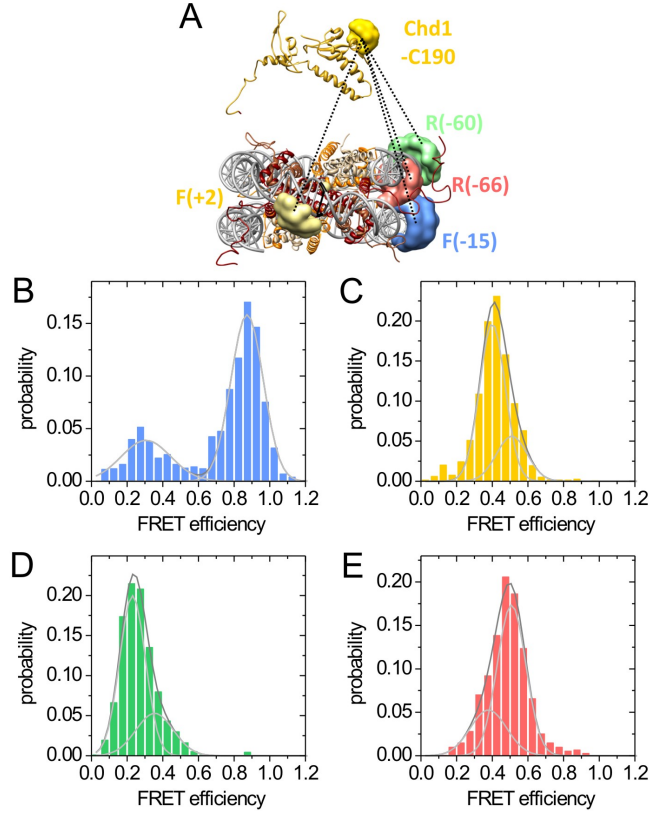


Figure A.3.3: Positioning of tandem chromodomains of Chd1 relative to the nucleosome core particle. (A) Schematic illustration of the set of smFRET measurements from all satellites to antenna Chd1-C190 on the tandem chromodomains. The images of nucleosome core particle and Chd1 chromodomains containing the dye position priors were prepared in Chimera [151]. The measurements are indicated by dotted black lines and the respective FRET efficiency histograms are shown in the color of the corresponding satellite (B-E)) (light blue for F(-15), yellow for F(+2), light green for R(-60), light red for R(-66)). All FRET efficiency histograms were best described by two Gaussians since two populations were visible, one major population comprising 72-75 % of the data and a subpopulation comprising 25-28 % of the data. (light gray: individual fits, dark gray: combined fits). The maximum FRET efficiencies of all fits are presented in Table A.3.1.

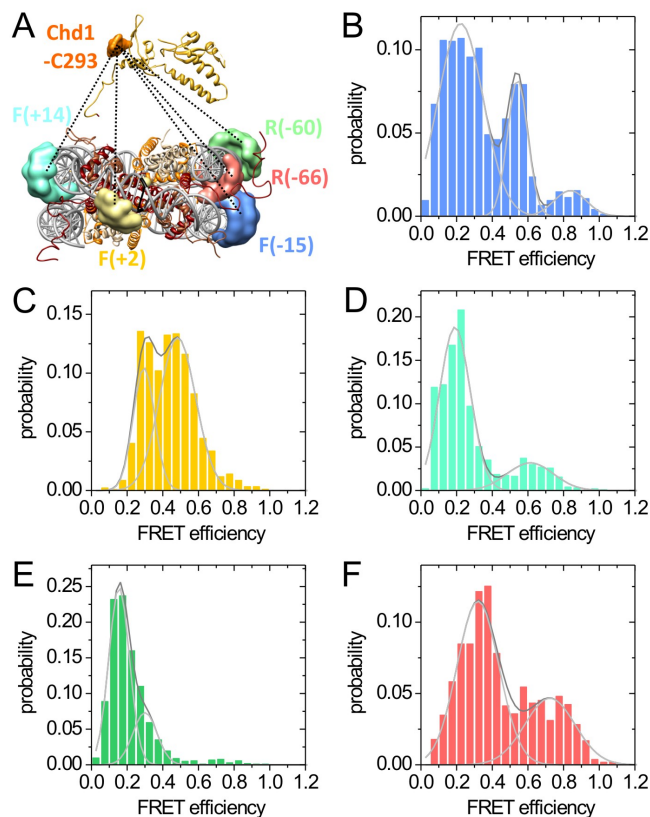


Figure A.3.4.: Positioning of tandem chromodomains of Chd1 relative to the nucleosome core particle. (A) Schematic illustration of the set of smFRET measurements from all satellites to antenna Chd1-C293 on the tandem chromodomains. The images of nucleosome core particle and Chd1 chromodomains containing the dye position priors were prepared in Chimera [151]. The measurements are indicated by dotted black lines and the respective FRET efficiency histograms are shown in the color of the corresponding satellite (B-F) (light blue for F(-15), yellow for F(+2), aquamarine for F(+14), light green for R(-60), light red for R(-66)). FRET efficiency histograms were best described with two Gaussians (three Gaussians for satellite F(-15)) since two (three) populations were visible, one major population comprising 68-79 % of the data and one (two) subpopulation comprising 21-32 % of the data (light gray: individual fits, dark gray: combined fits). The maximum FRET efficiencies of all fits are presented in Table A.3.1.

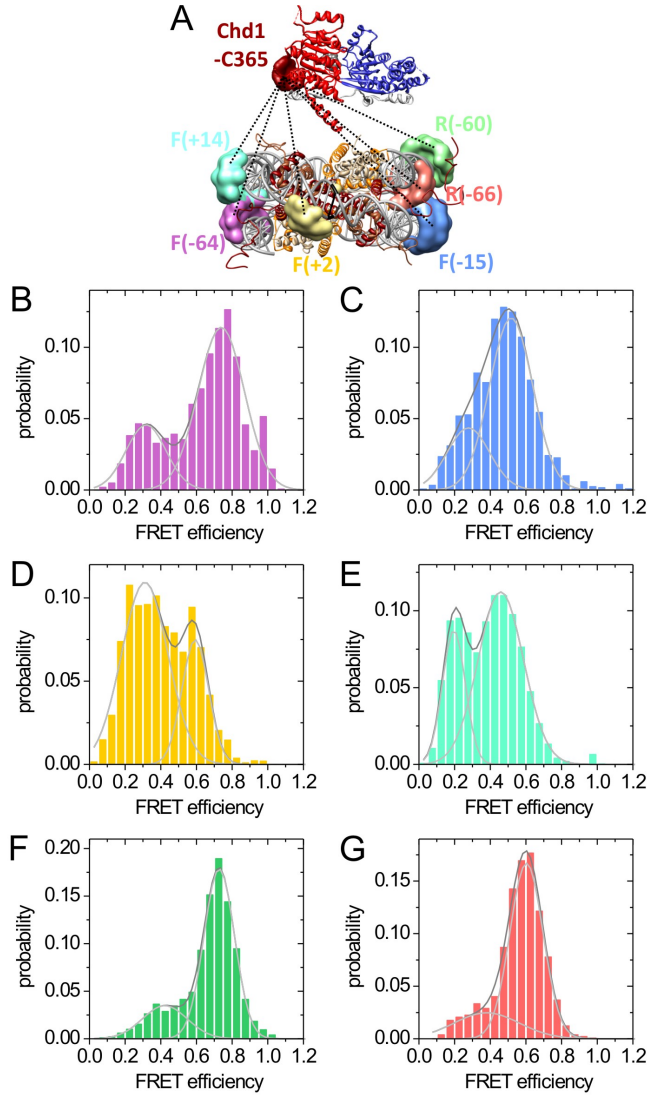


Figure A.3.5.: Positioning of ATPase lobe 1 of Chd1 relative to the nucleosome core particle. (A) Schematic illustration of the set of smFRET measurements from all satellites to antenna Chd1-C365 on the Snf2 ATPase domain. The images of nucleosome core particle and Chd1 ATPase domain containing the dye position priors were prepared in Chimera [151]. The measurements are indicated by dotted black lines and the respective FRET efficiency histograms are shown in the color of the corresponding satellite (B-G) (magenta for satellite F(-64), light blue for F(-15), yellow for F(+2), aquamarine for F(+14), light green for R(-60), light red for R(-66)). All FRET efficiency histograms were best described with two Gaussians since two populations were visible, one major population comprising 72-78 % of the data and a subpopulation comprising 22-28 % of the data (light gray: individual fits, dark gray: combined fits). The maximum FRET efficiencies of all fits are presented in Table A.3.1.

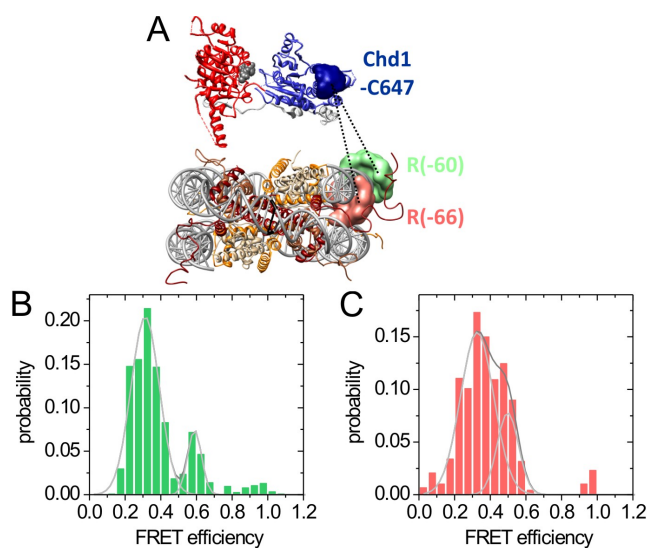


Figure A.3.6.: Positioning of ATPase lobe 2 of Chd1 relative to the nucleosome core particle. (A) Schematic illustration of the set of smFRET measurements from all satellites to antenna Chd1-C647 on the Snf2 ATPase domain. The images of nucleosome core particle and Chd1 ATPase domain containing the dye position priors were prepared in Chimera [151]. The measurements are indicated by dotted black lines and the respective FRET efficiency histograms are shown in the color of the corresponding satellite (B-C) (light green for R(-60), light red for R(-66)). The FRET efficiency histogram for satellite R(-60) was fit with two Gaussians, whereas the histogram for satellite R(-66) was best described by a single Gaussian. In general, measurements to antenna Chd1-C647 were tedious and yielded relatively low molecule numbers. (light gray: individual fits, dark gray: combined fits). The maximum FRET efficiencies of all fits are presented in Table A.3.1.

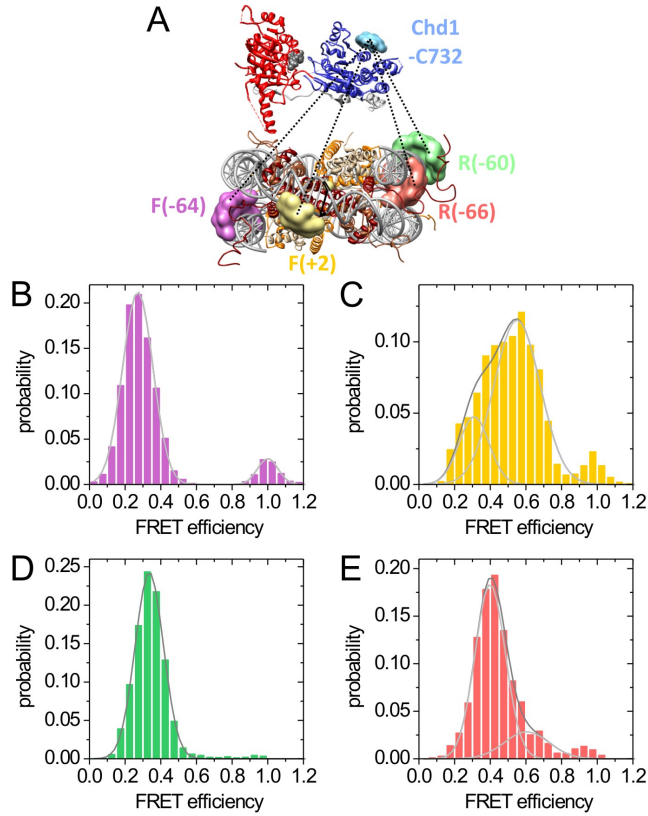


Figure A.3.7.: Positioning of ATPase lobe 2 of Chd1 relative to the nucleosome core particle. (A) Schematic illustration of the set of smFRET measurements from all satellites to antenna Chd1-C732 on the Snf2 ATPase domain. The images of nucleosome core particle and Chd1 ATPase domain containing the dye position priors were prepared in Chimera [151]. The measurements are indicated by dotted black lines and the respective FRET efficiency histograms are shown in the color of the corresponding satellite (B-E) (magenta for satellite F(-64), yellow for F(+2), light green for R(-60), light red for R(-66)). FRET efficiency histograms for satellites F(-64), F(+2) and R(-66) were best described with two Gaussians, one major population comprising 79-93 % of the data and a subpopulation comprising 21-7 % of the data (light gray: individual fits, dark gray: combined fits). The FRET distribution for satellite R(-60) was fit with one Gaussian. The maximum FRET efficiencies of all fits are presented in Table A.3.1.

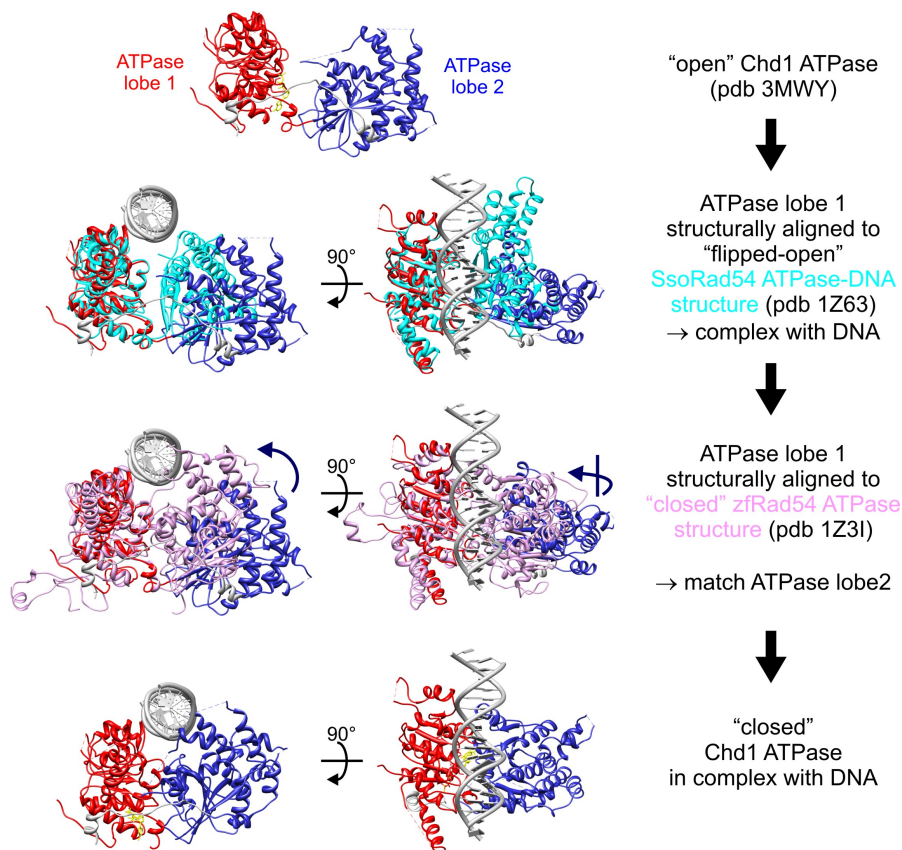


Figure A.3.8.: Structural modeling of a closed Chd1 ATPase domain in complex with DNA. First, ATPase lobe 1 of the open Chd1 ATPase structure (pdb ID: 3MWY, [147]) was structurally aligned with ATPase lobe 1 of the SsoRad54 ATPase-DNA complex structure (pdb ID: 1Z63, [159]) to obtain a model of the open Chd1 ATPase in complex with DNA. However, the ATPase in the SsoRad54 ATPase-DNA structure captures a different open form with ATPase lobe 2 flipped by 180° relative to ATPase lobe 1. Therefore in the next step, in order to obtain the closed ATPase conformation, Chd1 ATPase lobe 1 was structurally aligned to ATPase lobe 1 of the zebrafish Rad54 ATPase structure (pdb ID: 1Z3I), [260]), which shows a closed conformation, but in turn does not contain a DNA template. Finally, Chd1 ATPase lobe 2 was modeled to match ATPase lobe 2 of the closed zfATPase structure to yield a model of the closed Chd1 ATPase in complex with DNA.

Bibliography

- [1] F. Crick, *Central dogma of molecular biology.*, Nature **227**, 561 (1970).
- [2] S. Hahn, *Structure and mechanism of the RNA polymerase II transcription machinery.*, Nat Struct Mol Biol **11**, 394 (2004).
- [3] H. Boeger, D. A. Bushnell, R. Davis, J. Griesenbeck, Y. Lorch, J. S. Strattan, K. D. Westover, and R. D. Kornberg, *Structural basis of eukaryotic gene transcription.*, FEBS Lett **579**, 899 (2005).
- [4] R. G. Roeder, *The role of general initiation factors in transcription by RNA polymerase II.*, Trends Biochem Sci **21**, 327 (1996).
- [5] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, *Crystal structure of the nucleosome core particle at 2.8 Å resolution.*, Nature **389**, 251 (1997).
- [6] C. R. Clapier and B. R. Cairns, *The biology of chromatin remodeling complexes.*, Annu Rev Biochem **78**, 273 (2009).
- [7] B. Li, M. Carey, and J. L. Workman, *The role of chromatin during transcription.*, Cell **128**, 707 (2007).
- [8] T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, and S. Weiss, *Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor.*, Proc Natl Acad Sci U S A **93**, 6264 (1996).
- [9] C. Joo, H. Balci, Y. Ishitsuka, C. Buranachai, and T. Ha, *Advances in single-molecule fluorescence methods for molecular biology.*, Annu Rev Biochem **77**, 51 (2008).
- [10] A. N. Kapanidis, E. Margeat, S. O. Ho, E. Kortkhonjia, S. Weiss, and R. H. Ebright, *Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism.*, Science **314**, 1144 (2006).
- [11] A. Revyakin, C. Liu, R. H. Ebright, and T. R. Strick, *Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching.*, Science **314**, 1139 (2006).

Bibliography

- [12] H. Kim, G.-Q. Tang, S. S. Patel, and T. Ha, *Opening-closing dynamics of the mitochondrial transcription pre-initiation complex.*, Nucleic Acids Res (2011).
- [13] G. Li, M. Levitus, C. Bustamante, and J. Widom, *Rapid spontaneous accessibility of nucleosomal DNA.*, Nat Struct Mol Biol **12**, 46 (2005).
- [14] T. R. Blosser, J. G. Yang, M. D. Stone, G. J. Narlikar, and X. Zhuang, *Dynamics of nucleosome remodelling by individual ACF complexes.*, Nature **462**, 1022 (2009).
- [15] J. Andrecka, R. Lewis, F. Brückner, E. Lehmann, P. Cramer, and J. Michaelis, *Single-molecule tracking of mRNA exiting from RNA polymerase II.*, Proc Natl Acad Sci U S A **105**, 135 (2008).
- [16] J. L. Knight, V. Mekler, J. Mukhopadhyay, R. H. Ebright, and R. M. Levy, *Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: developing benchmarks of model quality and reliability.*, Biophys J **88**, 925 (2005).
- [17] M. Margittai, J. Widengren, E. Schweinberger, G. F. Schröder, S. Felekyan, E. Haustein, M. König, D. Fasshauer, H. Grubmüller, R. Jahn, and C. A. M. Seidel, *Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1.*, Proc Natl Acad Sci U S A **100**, 15516 (2003).
- [18] V. Mekler, E. Kortkhonjia, J. Mukhopadhyay, J. Knight, A. Revyakin, A. N. Kapandis, W. Niu, Y. W. Ebright, R. Levy, and R. H. Ebright, *Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex.*, Cell **108**, 599 (2002).
- [19] I. Rasnik, S. Myong, W. Cheng, T. M. Lohman, and T. Ha, *DNA-binding orientation and domain conformation of the E. coli rep helicase monomer bound to a partial duplex junction: single-molecule studies of fluorescently labeled enzymes.*, J Mol Biol **336**, 395 (2004).
- [20] A. Muschielok, J. Andrecka, A. Jawhari, F. Brückner, P. Cramer, and J. Michaelis, *A nano-positioning system for macromolecular structural analysis.*, Nat Methods **5**, 965 (2008).
- [21] J. Andrecka, B. Treutlein, M. A. I. Arcusa, A. Muschielok, R. Lewis, A. C. M. Cheung, P. Cramer, and J. Michaelis, *Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase II elongation complex.*, Nucleic Acids Res **37**, 5803 (2009).
- [22] A. Muschielok and J. Michaelis, *Application of the nano-positioning system to the analysis of fluorescence resonance energy transfer networks.*, J Phys Chem B **115**, 11927 (2011).

-
- [23] B. A. Hurwitz, J. and R. Diring, *Biochem. Biophys. Res. Commun.* **3**, 15 (1960).
- [24] S. B. Weiss, S. and T. Nakamoto, *Net synthesis of ribonucleic acid with a microbial enzyme requiring deoxyribonucleic acid and four ribonucleoside triphosphates.*, *J Biol Chem* **236**, PC18 (1961).
- [25] A. J. Herr, M. B. Jensen, T. Dalmay, and D. C. Baulcombe, *RNA polymerase IV directs silencing of endogenous DNA.*, *Science* **308**, 118 (2005).
- [26] T. Kanno, B. Huettel, M. F. Mette, W. Aufsatz, E. Jaligot, L. Daxinger, D. P. Kreil, M. Matzke, and A. J. M. Matzke, *Atypical RNA polymerase subunits required for RNA-directed DNA methylation.*, *Nat Genet* **37**, 761 (2005).
- [27] P. Cramer, K.-J. Armache, S. Baumli, S. Benkert, F. Brueckner, C. Buchen, G. E. Damsma, S. Dengl, S. R. Geiger, A. J. Jasiak, A. Jawhari, S. Jennebach, T. Kaminski, H. Kettenberger, C.-D. Kuhn, E. Lehmann, K. Leike, J. F. Sydow, and A. Vannini, *Structure of eukaryotic RNA polymerases.*, *Annu Rev Biophys* **37**, 337 (2008).
- [28] J. Q. Svejstrup, *The RNA polymerase II transcription cycle: cycling through chromatin.*, *Biochim Biophys Acta* **1677**, 64 (2004).
- [29] T. W. Sikorski and S. Buratowski, *The basal initiation machinery: beyond the general transcription factors.*, *Curr Opin Cell Biol* **21**, 344 (2009).
- [30] B. J. Venters, S. Wachi, T. N. Mavrich, B. E. Andersen, P. Jena, A. J. Sinnamon, P. Jain, N. S. Roller, C. Jiang, C. Hemeryck-Walsh, and B. F. Pugh, *A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces.*, *Mol Cell* **41**, 480 (2011).
- [31] B. R. Cairns, *The logic of chromatin architecture and remodelling at promoters.*, *Nature* **461**, 193 (2009).
- [32] M. Killeen, B. Coulombe, and J. Greenblatt, *Recombinant TBP, transcription factor IIB, and RAP30 are sufficient for promoter recognition by mammalian RNA polymerase II.*, *J Biol Chem* **267**, 9463 (1992).
- [33] G. Pan and J. Greenblatt, *Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site.*, *J Biol Chem* **269**, 30101 (1994).
- [34] M. Pal, A. S. Ponticelli, and D. S. Luse, *The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II.*, *Mol Cell* **19**, 101 (2005).
- [35] K. Adelman, M. T. Marr, J. Werner, A. Saunders, Z. Ni, E. D. Andrulis, and J. T. Lis, *Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS.*, *Mol Cell* **17**, 103 (2005).

Bibliography

- [36] S. Buratowski, *Progression through the RNA polymerase II CTD cycle.*, Mol Cell **36**, 541 (2009).
- [37] A. Meinhart, T. Kamenski, S. Hoepfner, S. Baumli, and P. Cramer, *A structural perspective of CTD function.*, Genes Dev **19**, 1401 (2005).
- [38] T. Owen-Hughes and T. Gkikopoulos, *Making sense of transcribing chromatin.*, Curr Opin Cell Biol (2012).
- [39] P. Richard and J. L. Manley, *Transcription termination by nuclear RNA polymerases.*, Genes Dev **23**, 1247 (2009).
- [40] E. Wahle and W. Keller, *The biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors.*, Annu Rev Biochem **61**, 419 (1992).
- [41] N. Yudkovsky, J. A. Ranish, and S. Hahn, *A transcription reinitiation intermediate that is stabilized by activator.*, Nature **408**, 225 (2000).
- [42] D. Kostrewa, M. E. Zeller, K.-J. Armache, M. Seizl, K. Leike, M. Thomm, and P. Cramer, *RNA polymerase II-TFIIB structure and mechanism of transcription initiation.*, Nature **462**, 323 (2009).
- [43] H. Kettenberger, K.-J. Armache, and P. Cramer, *Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage.*, Cell **114**, 347 (2003).
- [44] A. C. M. Cheung and P. Cramer, *Structural basis of RNA polymerase II backtracking, arrest and reactivation.*, Nature **471**, 249 (2011).
- [45] D. A. Bushnell, K. D. Westover, R. E. Davis, and R. D. Kornberg, *Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms.*, Science **303**, 983 (2004).
- [46] X. Liu, D. A. Bushnell, D. Wang, G. Calero, and R. D. Kornberg, *Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism.*, Science **327**, 206 (2010).
- [47] S. A. Darst, A. M. Edwards, E. W. Kubalek, and R. D. Kornberg, *Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution.*, Cell **66**, 121 (1991).
- [48] P. Cramer, D. A. Bushnell, and R. D. Kornberg, *Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution.*, Science **292**, 1863 (2001).
- [49] A. L. Gnatt, P. Cramer, J. Fu, D. A. Bushnell, and R. D. Kornberg, *Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution.*, Science **292**, 1876 (2001).

-
- [50] K.-J. Armache, H. Kettenberger, and P. Cramer, *Architecture of initiation-competent 12-subunit RNA polymerase II.*, Proc Natl Acad Sci U S A **100**, 6964 (2003).
- [51] H. Kettenberger, K.-J. Armache, and P. Cramer, *Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS.*, Mol Cell **16**, 955 (2004).
- [52] A. M. Edwards, C. M. Kane, R. A. Young, and R. D. Kornberg, *Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro.*, J Biol Chem **266**, 71 (1991).
- [53] F. J. Asturias, G. D. Meredith, C. L. Poglitsch, and R. D. Kornberg, *Two conformations of RNA polymerase II revealed by electron crystallography.*, J Mol Biol **272**, 536 (1997).
- [54] H.-T. Chen and S. Hahn, *Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex.*, Mol Cell **12**, 437 (2003).
- [55] D. B. Nikolov, H. Chen, E. D. Halay, A. A. Usheva, K. Hisatake, D. K. Lee, R. G. Roeder, and S. K. Burley, *Crystal structure of a TFIIB-TBP-TATA-element ternary complex.*, Nature **377**, 119 (1995).
- [56] E. J. Cho and S. Buratowski, *Evidence that transcription factor IIB is required for a post-assembly step in transcription initiation.*, J Biol Chem **274**, 25807 (1999).
- [57] T. S. Pardee, C. S. Bangur, and A. S. Ponticelli, *The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection.*, J Biol Chem **273**, 17859 (1998).
- [58] J. A. Ranish, N. Yudkovsky, and S. Hahn, *Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB.*, Genes Dev **13**, 49 (1999).
- [59] B. D. Dynlacht, T. Hoey, and R. Tjian, *Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation.*, Cell **66**, 563 (1991).
- [60] L. Tora, *A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription.*, Genes Dev **16**, 673 (2002).
- [61] P. Carninci et al., *Genome-wide analysis of mammalian promoter architecture and evolution.*, Nat Genet **38**, 626 (2006).

Bibliography

- [62] J. Ponjavic, B. Lenhard, C. Kai, J. Kawai, P. Carninci, Y. Hayashizaki, and A. Sandelin, *Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters.*, Genome Biol **7**, R78 (2006).
- [63] Y. Kim, J. H. Geiger, S. Hahn, and P. B. Sigler, *Crystal structure of a yeast TBP/TATA-box complex.*, Nature **365**, 512 (1993).
- [64] Z. F. Burton, M. Killeen, M. Sopta, L. G. Ortolan, and J. Greenblatt, *RAP30/74: a general initiation factor that binds to RNA polymerase II.*, Mol Cell Biol **8**, 1602 (1988).
- [65] M. A. Freire-Picos, S. Krishnamurthy, Z.-W. Sun, and M. Hampsey, *Evidence that the Tfg1/Tfg2 dimer interface of TFIIF lies near the active center of the RNA polymerase II initiation complex.*, Nucleic Acids Res **33**, 5045 (2005).
- [66] M. A. Ghazy, S. A. Brodie, M. L. Ammerman, L. M. Ziegler, and A. S. Ponticelli, *Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II.*, Mol Cell Biol **24**, 10975 (2004).
- [67] I. Pinto, W. H. Wu, J. G. Na, and M. Hampsey, *Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast.*, J Biol Chem **269**, 30569 (1994).
- [68] P. G. Rani, J. A. Ranish, and S. Hahn, *RNA polymerase II (Pol II)-TFIIF and Pol II-mediator complexes: the major stable Pol II complexes and their activity in transcription initiation and reinitiation.*, Mol Cell Biol **24**, 1709 (2004).
- [69] M. T. Killeen and J. F. Greenblatt, *The general transcription factor RAP30 binds to RNA polymerase II and prevents it from binding nonspecifically to DNA.*, Mol Cell Biol **12**, 30 (1992).
- [70] Z. A. Chen, A. Jawhari, L. Fischer, C. Buchen, S. Tahir, T. Kamenski, M. Ras-mussen, L. Lariviere, J.-C. Bukowski-Wills, M. Nilges, P. Cramer, and J. Rappsilber, *Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry.*, EMBO J **29**, 717 (2010).
- [71] J. Fishburn and S. Hahn, *Architecture of the yeast RNA polymerase II open complex and regulation of activity by TFIIF.*, Mol Cell Biol **32**, 12 (2012).
- [72] M. H. Larson, R. Landick, and S. M. Block, *Single-molecule studies of RNA poly-merase: one singular sensation, every little step it takes.*, Mol Cell **41**, 249 (2011).
- [73] C. Bustamante, W. Cheng, Y. X. Mejia, and Y. X. Mejia, *Revisiting the central dogma one molecule at a time.*, Cell **144**, 480 (2011).

-
- [74] D. G. Vassylyev, S.-i. Sekine, O. Laptenko, J. Lee, M. N. Vassylyeva, S. Borukhov, and S. Yokoyama, *Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution.*, Nature **417**, 712 (2002).
- [75] D. G. Vassylyev, M. N. Vassylyeva, A. Perederina, T. H. Tahirov, and I. Artsimovitch, *Structural basis for transcription elongation by bacterial RNA polymerase.*, Nature **448**, 157 (2007).
- [76] C. Rivetti, M. Guthold, and C. Bustamante, *Wrapping of DNA around the E.coli RNA polymerase open promoter complex.*, EMBO J **18**, 4464 (1999).
- [77] A. N. Kapanidis, E. Margeat, T. A. Laurence, S. Doose, S. O. Ho, J. Mukhopadhyay, E. Kortkhonjia, V. Mekler, R. H. Ebright, and S. Weiss, *Retention of transcription initiation factor sigma70 in transcription elongation: single-molecule analysis.*, Mol Cell **20**, 347 (2005).
- [78] E. Margeat, A. N. Kapanidis, P. Tinnefeld, Y. Wang, J. Mukhopadhyay, R. H. Ebright, and S. Weiss, *Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes.*, Biophys J **90**, 1419 (2006).
- [79] M. D. Wang, M. J. Schnitzer, H. Yin, R. Landick, J. Gelles, and S. M. Block, *Force and velocity measured for single molecules of RNA polymerase.*, Science **282**, 902 (1998).
- [80] R. J. Davenport, G. J. Wuite, R. Landick, and C. Bustamante, *Single-molecule study of transcriptional pausing and arrest by E. coli RNA polymerase.*, Science **287**, 2497 (2000).
- [81] K. Adelman, A. La Porta, T. J. Santangelo, J. T. Lis, J. W. Roberts, and M. D. Wang, *Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior.*, Proc Natl Acad Sci U S A **99**, 13538 (2002).
- [82] N. R. Forde, D. Izhaky, G. R. Woodcock, G. J. L. Wuite, and C. Bustamante, *Using mechanical force to probe the mechanism of pausing and arrest during continuous elongation by Escherichia coli RNA polymerase.*, Proc Natl Acad Sci U S A **99**, 11682 (2002).
- [83] K. M. Herbert, A. La Porta, B. J. Wong, R. A. Mooney, K. C. Neuman, R. Landick, and S. M. Block, *Sequence-resolved detection of pausing by single RNA polymerase molecules.*, Cell **125**, 1083 (2006).
- [84] J. W. Shaevitz, E. A. Abbondanzieri, R. Landick, and S. M. Block, *Backtracking by single RNA polymerase molecules observed at near-base-pair resolution.*, Nature **426**, 684 (2003).

Bibliography

- [85] Y. X. Mejia, H. Mao, N. R. Forde, and C. Bustamante, *Thermal probing of E. coli RNA polymerase off-pathway mechanisms.*, J Mol Biol **382**, 628 (2008).
- [86] E. A. Galburt, S. W. Grill, A. Wiedmann, L. Lubkowska, J. Choy, E. Nogales, M. Kashlev, and C. Bustamante, *Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner.*, Nature **446**, 820 (2007).
- [87] R. V. Dalal, M. H. Larson, K. C. Neuman, J. Gelles, R. Landick, and S. M. Block, *Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing.*, Mol Cell **23**, 231 (2006).
- [88] M. H. Larson, W. J. Greenleaf, R. Landick, and S. M. Block, *Applied force reveals mechanistic and energetic details of transcription termination.*, Cell **132**, 971 (2008).
- [89] H. Yin, I. Artsimovitch, R. Landick, and J. Gelles, *Nonequilibrium mechanism of transcription termination from observations of single RNA polymerase molecules.*, Proc Natl Acad Sci U S A **96**, 13124 (1999).
- [90] C. Hodges, L. Bintu, L. Lubkowska, M. Kashlev, and C. Bustamante, *Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II.*, Science **325**, 626 (2009).
- [91] L. Bintu, M. Kopaczynska, C. Hodges, L. Lubkowska, M. Kashlev, and C. Bustamante, *The elongation rate of RNA polymerase determines the fate of transcribed nucleosomes.*, Nat Struct Mol Biol **18**, 1394 (2011).
- [92] J. Jin, L. Bai, D. S. Johnson, R. M. Fulbright, M. L. Kireeva, M. Kashlev, and M. D. Wang, *Synergistic action of RNA polymerases in overcoming the nucleosomal barrier.*, Nat Struct Mol Biol **17**, 745 (2010).
- [93] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, *Real-time kinetics of gene activity in individual bacteria.*, Cell **123**, 1025 (2005).
- [94] D. R. Larson, R. H. Singer, and D. Zenklusen, *A single molecule view of gene expression.*, Trends Cell Biol **19**, 630 (2009).
- [95] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, *Stochastic mRNA synthesis in mammalian cells.*, PLoS Biol **4**, e309 (2006).
- [96] D. Zenklusen, D. R. Larson, and R. H. Singer, *Single-RNA counting reveals alternative modes of gene expression in yeast.*, Nat Struct Mol Biol **15**, 1263 (2008).
- [97] D. R. Larson, D. Zenklusen, B. Wu, J. A. Chao, and R. H. Singer, *Real-time observation of transcription initiation and elongation on an endogenous yeast gene.*, Science **332**, 475 (2011).

-
- [98] A. L. Olins and D. E. Olins, *Spheroid chromatin units (v bodies).*, Science **183**, 330 (1974).
- [99] C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder, and T. J. Richmond, *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution.*, J Mol Biol **319**, 1097 (2002).
- [100] E. I. Campos and D. Reinberg, *Histones: annotating chromatin.*, Annu Rev Genet **43**, 559 (2009).
- [101] D. J. Tremethick, *Higher-order structures of chromatin: the elusive 30 nm fiber.*, Cell **128**, 651 (2007).
- [102] G. Orphanides and D. Reinberg, *RNA polymerase II elongation through chromatin.*, Nature **407**, 471 (2000).
- [103] P. J. Laybourn and J. T. Kadonaga, *Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II.*, Science **254**, 238 (1991).
- [104] D. S. Luse, L. C. Spangler, and A. Újvári, *Efficient and rapid nucleosome traversal by RNA polymerase II depends on a combination of transcript elongation factors.*, J Biol Chem **286**, 6040 (2011).
- [105] V. M. Studitsky, G. A. Kassavetis, E. P. Geiduschek, and G. Felsenfeld, *Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase.*, Science **278**, 1960 (1997).
- [106] J. Widom, *Equilibrium and dynamic nucleosome stability.*, Methods Mol Biol **119**, 61 (1999).
- [107] J. Widom, *Role of DNA sequence in nucleosome stability and dynamics.*, Q Rev Biophys **34**, 269 (2001).
- [108] N. Kaplan, I. K. Moore, Y. Fondufe-Mittendorf, A. J. Gossett, D. Tillo, Y. Field, E. M. LeProust, T. R. Hughes, J. D. Lieb, J. Widom, and E. Segal, *The DNA-encoded nucleosome organization of a eukaryotic genome.*, Nature **458**, 362 (2009).
- [109] E. Segal, Y. Fondufe-Mittendorf, L. Chen, A. Thåström, Y. Field, I. K. Moore, J.-P. Z. Wang, and J. Widom, *A genomic code for nucleosome positioning.*, Nature **442**, 772 (2006).
- [110] M.-L. Visnapuu and E. C. Greene, *Single-molecule imaging of DNA curtains reveals intrinsic energy landscapes for nucleosome deposition.*, Nat Struct Mol Biol **16**, 1056 (2009).

Bibliography

- [111] P. T. Lowary and J. Widom, *New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning.*, J Mol Biol **276**, 19 (1998).
- [112] M. A. Hall, A. Shundrovsky, L. Bai, R. M. Fulbright, J. T. Lis, and M. D. Wang, *High-resolution dynamic mapping of histone-DNA interactions in a nucleosome.*, Nat Struct Mol Biol **16**, 124 (2009).
- [113] S. Mihardja, A. J. Spakowitz, Y. Zhang, and C. Bustamante, *Effect of force on mononucleosomal dynamics.*, Proc Natl Acad Sci U S A **103**, 15871 (2006).
- [114] J. D. Anderson, A. Thåström, and J. Widom, *Spontaneous access of proteins to buried nucleosomal DNA target sites occurs via a mechanism that is distinct from nucleosome translocation.*, Mol Cell Biol **22**, 7147 (2002).
- [115] W. J. A. Koopmans, A. Brehm, C. Logie, T. Schmidt, and J. van Noort, *Single-pair FRET microscopy reveals mononucleosome dynamics.*, J Fluoresc **17**, 785 (2007).
- [116] H. Neumann, S. M. Hancock, R. Buning, A. Routh, L. Chapman, J. Somers, T. Owen-Hughes, J. van Noort, D. Rhodes, and J. W. Chin, *A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation.*, Mol Cell **36**, 153 (2009).
- [117] A. J. Andrews and K. Luger, *Nucleosome structure(s) and stability: variations on a theme.*, Annu Rev Biophys **40**, 99 (2011).
- [118] V. Böhm, A. R. Hieb, A. J. Andrews, A. Gansen, A. Rocker, K. Tóth, K. Luger, and J. Langowski, *Nucleosome accessibility governed by the dimer/tetramer interface.*, Nucleic Acids Res **39**, 3093 (2011).
- [119] M. E. Fairman-Williams, U.-P. Guenther, and E. Jankowsky, *SF1 and SF2 helicases: family matters.*, Curr Opin Struct Biol **20**, 313 (2010).
- [120] C. G. A. Marfella and A. N. Imbalzano, *The Chd family of chromatin remodelers.*, Mutat Res **618**, 30 (2007).
- [121] T. Woodage, M. A. Basrai, A. D. Baxevanis, P. Hieter, and F. S. Collins, *Characterization of the CHD family of proteins.*, Proc Natl Acad Sci U S A **94**, 11472 (1997).
- [122] R. Paro and D. S. Hogness, *The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of Drosophila.*, Proc Natl Acad Sci U S A **88**, 263 (1991).
- [123] D. G. Stokes, K. D. Tartof, and R. P. Perry, *CHD1 is concentrated in interbands and puffed regions of Drosophila polytene chromosomes.*, Proc Natl Acad Sci U S A **93**, 7137 (1996).

-
- [124] V. Delmas, D. G. Stokes, and R. P. Perry, *A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain.*, Proc Natl Acad Sci U S A **90**, 2414 (1993).
- [125] H. G. Tran, D. J. Steger, V. R. Iyer, and A. D. Johnson, *The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor.*, EMBO J **19**, 2323 (2000).
- [126] V. K. Gangaraju and B. Bartholomew, *Mechanisms of ATP dependent chromatin remodeling.*, Mutat Res **618**, 3 (2007).
- [127] A. Lusser, D. L. Urwin, and J. T. Kadonaga, *Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly.*, Nat Struct Mol Biol **12**, 160 (2005).
- [128] C. Stockdale, A. Flaus, H. Ferreira, and T. Owen-Hughes, *Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes.*, J Biol Chem **281**, 16279 (2006).
- [129] T. Ito, M. Bulger, M. J. Pazin, R. Kobayashi, and J. T. Kadonaga, *ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor.*, Cell **90**, 145 (1997).
- [130] D. F. Corona, G. Längst, C. R. Clapier, E. J. Bonte, S. Ferrari, J. W. Tamkun, and P. B. Becker, *ISWI is an ATP-dependent nucleosome remodeling factor.*, Mol Cell **3**, 239 (1999).
- [131] M. E. Gelbart, T. Rechsteiner, T. J. Richmond, and T. Tsukiyama, *Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates.*, Mol Cell Biol **21**, 2098 (2001).
- [132] R. Simic, D. L. Lindstrom, H. G. Tran, K. L. Roinick, P. J. Costa, A. D. Johnson, G. A. Hartzog, and K. M. Arndt, *Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes.*, EMBO J **22**, 1846 (2003).
- [133] D. E. Kelley, D. G. Stokes, and R. P. Perry, *CHD1 interacts with SSRP1 and depends on both its chromodomain and its ATPase/helicase-like domain for proper association with chromatin.*, Chromosoma **108**, 10 (1999).
- [134] N. J. Krogan, M. Kim, S. H. Ahn, G. Zhong, M. S. Kobor, G. Cagney, A. Emili, A. Shilatifard, S. Buratowski, and J. F. Greenblatt, *RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach.*, Mol Cell Biol **22**, 6979 (2002).

Bibliography

- [135] C. Alén, N. A. Kent, H. S. Jones, J. O’Sullivan, A. Aranda, and N. J. Proudfoot, *A role for chromatin remodeling in transcriptional termination by RNA polymerase II.*, Mol Cell **10**, 1441 (2002).
- [136] J. J. Lin, L. W. Lehmann, G. Bonora, R. Sridharan, A. A. Vashisht, N. Tran, K. Plath, J. A. Wohlschlegel, and M. Carey, *Mediator coordinates PIC assembly with recruitment of CHD1.*, Genes Dev **25**, 2198 (2011).
- [137] T. Tsukiyama, J. Palmer, C. C. Landel, J. Shiloach, and C. Wu, *Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae.*, Genes Dev **13**, 686 (1999).
- [138] A. Y. Konev, M. Tribus, S. Y. Park, V. Podhraski, C. Y. Lim, A. V. Emelyanov, E. Vershilova, V. Pirrotta, J. T. Kadonaga, A. Lusser, and D. V. Fyodorov, *CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo.*, Science **317**, 1087 (2007).
- [139] J. Walfridsson, O. Khorosjutina, P. Matikainen, C. M. Gustafsson, and K. Ekwall, *A genome-wide role for CHD remodelling factors and Nap1 in nucleosome disassembly.*, EMBO J **26**, 2868 (2007).
- [140] A. Gaspar-Maia, A. Alajem, F. Polesso, R. Sridharan, M. J. Mason, A. Heidersbach, J. Ramalho-Santos, M. T. McManus, K. Plath, E. Meshorer, and M. Ramalho-Santos, *Chd1 regulates open chromatin and pluripotency of embryonic stem cells.*, Nature **460**, 863 (2009).
- [141] D. G. Stokes and R. P. Perry, *DNA-binding and chromatin localization properties of CHD1.*, Mol Cell Biol **15**, 2745 (1995).
- [142] J. F. Flanagan, B. J. Blus, D. Kim, K. L. Clines, F. Rastinejad, and S. Khorasanizadeh, *Molecular implications of evolutionary differences in CHD double chromodomains.*, J Mol Biol **369**, 334 (2007).
- [143] R. J. Sims, 3rd, C.-F. Chen, H. Santos-Rosa, T. Kouzarides, S. S. Patel, and D. Reinberg, *Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains.*, J Biol Chem **280**, 41789 (2005).
- [144] M. Okuda, M. Horikoshi, and Y. Nishimura, *Structural polymorphism of chromodomains in Chd1.*, J Mol Biol **365**, 1047 (2007).
- [145] S. Morettini, M. Tribus, A. Zeilner, J. Sebald, B. Campo-Fernandez, G. Scheran, H. Wörle, V. Podhraski, D. V. Fyodorov, and A. Lusser, *The chromodomains of CHD1 are critical for enzymatic activity but less important for chromatin localization.*, Nucleic Acids Res **39**, 3103 (2011).

- [146] J. F. Flanagan, L.-Z. Mi, M. Chruszcz, M. Cymborowski, K. L. Clines, Y. Kim, W. Minor, F. Rastinejad, and S. Khorasanizadeh, *Double chromodomains cooperate to recognize the methylated histone H3 tail.*, Nature **438**, 1181 (2005).
- [147] G. Hauk, J. N. McKnight, I. M. Nodelman, and G. D. Bowman, *The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor.*, Mol Cell **39**, 711 (2010).
- [148] I. Forné, J. Ludwigsen, A. Imhof, P. B. Becker, and F. Mueller-Planitz, *Probing the conformation of the ISWI ATPase domain with genetically encoded photoreactive crosslinkers and mass spectrometry.*, Mol Cell Proteomics **11**, M111.012088 (2012).
- [149] D. P. Ryan, R. Sundaramoorthy, D. Martin, V. Singh, and T. Owen-Hughes, *The DNA-binding domain of the Chd1 chromatin-remodelling enzyme contains SANT and SLIDE domains.*, EMBO J **30**, 2596 (2011).
- [150] A. Sharma, K. R. Jenkins, A. Héroux, and G. D. Bowman, *Crystal structure of the chromodomain helicase DNA-binding protein 1 (Chd1) DNA-binding domain in complex with DNA.*, J Biol Chem **286**, 42099 (2011).
- [151] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, *UCSF Chimera—a visualization system for exploratory research and analysis.*, J Comput Chem **25**, 1605 (2004).
- [152] J. N. McKnight, K. R. Jenkins, I. M. Nodelman, T. Escobar, and G. D. Bowman, *Extranucleosomal DNA binding directs nucleosome sliding by Chd1.*, Mol Cell Biol **31**, 4746 (2011).
- [153] T. Grüne, J. Brzeski, A. Eberharter, C. R. Clapier, D. F. V. Corona, P. B. Becker, and C. W. Müller, *Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI.*, Mol Cell **12**, 449 (2003).
- [154] K. Yamada, T. D. Frouws, B. Angst, D. J. Fitzgerald, C. DeLuca, K. Schimmele, D. F. Sargent, and T. J. Richmond, *Structure and mechanism of the chromatin remodelling factor ISW1a.*, Nature **472**, 448 (2011).
- [155] A. Patel, J. N. McKnight, P. Genzor, and G. D. Bowman, *Identification of residues in chromodomain helicase DNA-binding protein 1 (Chd1) required for coupling ATP hydrolysis to nucleosome sliding.*, J Biol Chem **286**, 43984 (2011).
- [156] J. M. Gottesfeld and K. Luger, *Energetics and affinity of the histone octamer for defined DNA sequences.*, Biochemistry **40**, 10927 (2001).
- [157] A. Flaus and T. Owen-Hughes, *Dynamic properties of nucleosomes during thermal and ATP-driven mobilization.*, Mol Cell Biol **23**, 7767 (2003).

Bibliography

- [158] T. J. Richmond and C. A. Davey, *The structure of DNA in the nucleosome core.*, Nature **423**, 145 (2003).
- [159] H. Dürr, C. Körner, M. Müller, V. Hickmann, and K.-P. Hopfner, *X-ray structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA.*, Cell **121**, 363 (2005).
- [160] S. Dumont, W. Cheng, V. Serebrov, R. K. Beran, I. Tinoco, Jr, A. M. Pyle, and C. Bustamante, *RNA translocation and unwinding mechanism of HCV NS3 helicase and its coordination by ATP.*, Nature **439**, 105 (2006).
- [161] T. D. Yager and K. E. van Holde, *Dynamics and equilibria of nucleosomes at elevated ionic strength.*, J Biol Chem **259**, 4212 (1984).
- [162] I. M. Kulic and H. Schiessel, *Chromatin dynamics: nucleosomes go mobile through twist defects.*, Phys Rev Lett **91**, 148103 (2003).
- [163] G. Längst and P. B. Becker, *Nucleosome remodeling: one mechanism, many phenomena?*, Biochim Biophys Acta **1677**, 58 (2004).
- [164] M. N. Kagalwala, B. J. Glaus, W. Dang, M. Zofall, and B. Bartholomew, *Topography of the ISW2-nucleosome complex: insights into nucleosome spacing and chromatin remodeling.*, EMBO J **23**, 2092 (2004).
- [165] M. Zofall, J. Persinger, S. R. Kassabov, and B. Bartholomew, *Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome.*, Nat Struct Mol Biol **13**, 339 (2006).
- [166] R. Schwanbeck, H. Xiao, and C. Wu, *Spatial contacts and nucleosome step movements induced by the NURF chromatin remodeling complex.*, J Biol Chem **279**, 39933 (2004).
- [167] W. Dang and B. Bartholomew, *Domain architecture of the catalytic subunit in the ISW2-nucleosome complex.*, Mol Cell Biol **27**, 8306 (2007).
- [168] V. K. Gangaraju, P. Prasad, A. Srour, M. N. Kagalwala, and B. Bartholomew, *Conformational changes associated with template commitment in ATP-dependent chromatin remodeling by ISW2.*, Mol Cell **35**, 58 (2009).
- [169] A. Saha, J. Wittmeyer, and B. R. Cairns, *Chromatin remodeling through directional DNA translocation from an internal nucleosomal site.*, Nat Struct Mol Biol **12**, 747 (2005).
- [170] R. Strohner, M. Wachsmuth, K. Dachauer, J. Mazurkiewicz, J. Hochstatter, K. Rippe, and G. Längst, *A 'loop recapture' mechanism for ACF-dependent nucleosome remodeling.*, Nat Struct Mol Biol **12**, 683 (2005).

-
- [171] K.-P. Hopfner and J. Michaelis, *Mechanisms of nucleic acid translocases: lessons from structural biology and single-molecule biophysics.*, Curr Opin Struct Biol **17**, 87 (2007).
- [172] L. R. Racki, J. G. Yang, N. Naber, P. D. Partensky, A. Acevedo, T. J. Purcell, R. Cooke, Y. Cheng, and G. J. Narlikar, *The chromatin remodeller ACF acts as a dimeric motor to space nucleosomes.*, Nature **462**, 1016 (2009).
- [173] B. R. Cairns, *Chromatin remodeling: insights and intrigue from single-molecule studies.*, Nat Struct Mol Biol **14**, 989 (2007).
- [174] J. L. Killian, M. Li, M. Y. Sheinin, and M. D. Wang, *Recent advances in single molecule studies of nucleosomes.*, Curr Opin Struct Biol **22**, 80 (2012).
- [175] Y. Zhang, C. L. Smith, A. Saha, S. W. Grill, S. Mihardja, S. B. Smith, B. R. Cairns, C. L. Peterson, and C. Bustamante, *DNA translocation and loop formation mechanism of chromatin remodeling by SWI/SNF and RSC.*, Mol Cell **24**, 559 (2006).
- [176] G. Lia, E. Praly, H. Ferreira, C. Stockdale, Y. C. Tse-Dinh, D. Dunlap, V. Croquette, D. Bensimon, and T. Owen-Hughes, *Direct observation of DNA distortion by the RSC complex.*, Mol Cell **21**, 417 (2006).
- [177] A. Shundrovsky, C. L. Smith, J. T. Lis, C. L. Peterson, and M. D. Wang, *Probing SWI/SNF remodeling of the nucleosome by unzipping single DNA molecules.*, Nat Struct Mol Biol **13**, 549 (2006).
- [178] J. Herschel, *On a case of superficial colour presented by a homogenous liquid internally colourless.*, Phil.Trans.R.Soc.Lond. **135**, 143 (1845).
- [179] G. G. Stokes, *On the change of regradability of light.*, Phil.Trans.R.Soc.Lond. **142**, 463 (1852).
- [180] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, 3 edition, 2006.
- [181] J. Vogelsang, R. Kasper, C. Steinhauer, B. Person, M. Heilemann, M. Sauer, and P. Tinnefeld, *A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes.*, Angew Chem Int Ed Engl **47**, 5465 (2008).
- [182] T. Förster, *Energiewanderung und Fluoreszenz.*, Naturwissenschaften **33(6)**, 166 (1946).
- [183] T. Förster, *Zwischenmolekulare Energiewanderung und Fluoreszenz.*, Annalen der Physik **437**, 55 (1948).
- [184] L. Stryer and R. P. Haugland, *Energy transfer: a spectroscopic ruler.*, Proc Natl Acad Sci U S A **58**, 719 (1967).

Bibliography

- [185] R. E. Dale, J. Eisinger, and W. E. Blumberg, *The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer.*, Biophys J **26**, 161 (1979).
- [186] J. Yin, P. D. Straight, S. M. McLoughlin, Z. Zhou, A. J. Lin, D. E. Golan, N. L. Kelleher, R. Kolter, and C. T. Walsh, *Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase.*, Proc Natl Acad Sci U S A **102**, 15815 (2005).
- [187] J. Yin, A. J. Lin, D. E. Golan, and C. T. Walsh, *Site-specific protein labeling by Sfp phosphopantetheinyl transferase.*, Nat Protoc **1**, 280 (2006).
- [188] X. Shi, Y. Jung, L.-J. Lin, C. Liu, C. Wu, I. K. O. Cann, and T. Ha, *Quantitative fluorescence labeling of aldehyde-tagged proteins for single-molecule imaging.*, Nat Methods (2012).
- [189] I. S. Carrico, B. L. Carlson, and C. R. Bertozzi, *Introducing genetically encoded aldehydes into proteins.*, Nat Chem Biol **3**, 321 (2007).
- [190] J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, and P. G. Schultz, *Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli.*, J Am Chem Soc **124**, 9026 (2002).
- [191] J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang, and P. G. Schultz, *An expanded eukaryotic genetic code.*, Science **301**, 964 (2003).
- [192] Moerner and Kador, *Optical detection and spectroscopy of single molecules in a solid.*, Phys Rev Lett **62**, 2535 (1989).
- [193] Orrit and Bernard, *Single pentacene molecules detected by fluorescence excitation in a p-terphenyl crystal.*, Phys Rev Lett **65**, 2716 (1990).
- [194] T. Funatsu, Y. Harada, M. Tokunaga, K. Saito, and T. Yanagida, *Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution.*, Nature **374**, 555 (1995).
- [195] D. Axelrod, N. L. Thompson, and T. P. Burghardt, *Total internal inflection fluorescent microscopy.*, J Microsc **129**, 19 (1983).
- [196] R. A. Lewis, *Untersuchung und Aufklärung des molekularen Mechanismus von Enzymen der SWI2/SNF2-Familie mit Hilfe der Einzelmolekül-Fluoreszenzmikroskopie.*, PhD thesis, Ludwig-Maximilians-Universität München, Fakultät Chemie und Pharmazie, 2009.
- [197] J. Hesse, M. Sonnleitner, A. Sonnleitner, G. Freudenthaler, J. Jacak, O. Höglinger, H. Schindler, and G. J. Schütz, *Single-molecule reader for high-throughput bioanalysis.*, Anal Chem **76**, 5960 (2004).

-
- [198] A. N. Kapanidis, T. A. Laurence, N. K. Lee, E. Margeat, X. Kong, and S. Weiss, *Alternating-laser excitation of single molecules.*, Acc Chem Res **38**, 523 (2005).
- [199] M. L. Bennink, S. H. Leuba, G. H. Leno, J. Zlatanova, B. G. de Grooth, and J. Greve, *Unfolding individual nucleosomes by stretching single chromatin fibers with optical tweezers.*, Nat Struct Biol **8**, 606 (2001).
- [200] J. Andrecka, *Single molecule fluorescence studies of the RNA polymerase II elongation complex.*, PhD thesis, Ludwig-Maximilians-Universität München, Fakultät Chemie und Pharmazie, 2009.
- [201] G. Heiss, *Single-molecule microscopy study of nano-systems: from synthetic photo-switchable nano-devices to the dynamics of naturally occurring transcription factors.*, PhD thesis, Ludwig-Maximilians-Universität München, Fakultät Chemie und Pharmazie, 2011.
- [202] L. Baum and T. Petrie, *Statistical inference for probabilistic functions of finite state markov chains.*, The annals of mathematical statistics **37**, 1554 (1966).
- [203] J. Raviv, *Decision making in markov chains applied to problem of pattern recognition.*, IEEE Transactions on Information Theory **13**, 536 (1967).
- [204] F. Jelinek, *Continuous speech recognition by statistical-methods.*, Proceedings of the IEEE **64**, 532 (1976).
- [205] L. Rabiner and B. Juang, *An introduction to hidden markov models.*, ASSP Magazine, IEEE **3**, 4 (1986).
- [206] G. A. Churchill, *Stochastic models for heterogeneous DNA sequences.*, Bull Math Biol **51**, 79 (1989).
- [207] R. Durbin, *Biological sequence analysis: Probabilistic models of proteins and nucleic acids.*, Cambridge University Press, 1998.
- [208] D. A. Coast, G. G. Cano, and S. A. Briller, *Use of hidden Markov models for electrocardiographic signal analysis.*, J Electrocardiol **23 Suppl**, 184 (1990).
- [209] F. G. Ball and J. A. Rice, *Stochastic models for ion channels: introduction and bibliography.*, Math Biosci **112**, 189 (1992).
- [210] M. Kruithof and J. van Noort, *Hidden Markov analysis of nucleosome unwrapping under force.*, Biophys J **96**, 3708 (2009).
- [211] T. C. Messina, H. Kim, J. T. Giurleo, and D. S. Talaga, *Hidden Markov model analysis of multichromophore photobleaching.*, J Phys Chem B **110**, 16366 (2006).

Bibliography

- [212] S. A. McKinney, C. Joo, and T. Ha, *Analysis of single-molecule FRET trajectories using hidden Markov modeling.*, Biophys J **91**, 1941 (2006).
- [213] C. Joo, S. A. McKinney, M. Nakamura, I. Rasnik, S. Myong, and T. Ha, *Real-time observation of RecA filament dynamics with single monomer resolution.*, Cell **126**, 515 (2006).
- [214] K. Murphy, *MATLAB HMM-Toolbox*, 1998.
- [215] A. Viterbi, *Error bounds for convolutional codes and an asymptotically optimum decoding algorithm.*, IEEE Transactions on Information Theory **13**, 260 (1967).
- [216] P. T. S. G. Baum, L. and N. Weiss, *A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains.*, The annals of mathematical statistics **41**, 164 (1970).
- [217] A. K. Wozniak, G. F. Schröder, H. Grubmüller, C. A. M. Seidel, and F. Oesterhelt, *Single-molecule FRET measures bends and kinks in DNA.*, Proc Natl Acad Sci U S A **105**, 18337 (2008).
- [218] J. Mukhopadhyay, A. N. Kapanidis, V. Mekler, E. Kortkhonjia, Y. W. Ebright, and R. H. Ebright, *Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA.*, Cell **106**, 453 (2001).
- [219] D. Sivis, *Data analysis: a Bayesian tutorial*, Oxford University Press, 2006.
- [220] D. Grohmann, J. Nagy, A. Chakraborty, D. Klose, D. Fielden, R. H. Ebright, J. Michaelis, and F. Werner, *The initiation factor tfe and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation.*, Mol Cell **43**, 263 (2011).
- [221] H. Balci, S. Arslan, S. Myong, T. M. Lohman, and T. Ha, *Single-molecule nanopositioning: structural transitions of a helicase-DNA complex during ATP hydrolysis.*, Biophys J **101**, 976 (2011).
- [222] A. M. Muschielok, *Development and application of a quantitative analysis method for fluorescence resonance energy transfer localization experiments.*, PhD thesis, Ludwig-Maximilians-Universität München, Fakultät Chemie und Pharmazie, 2010.
- [223] T. Bayes, *An essay towards solving a problem in the doctrine of chances.*, Philosophical Transactions of the Royal Society of London **53**, 330 (1763).
- [224] J. Skilling, *Nested Sampling for General Bayesian Computation.*, Bayesian Analysis **1**, 833 (2006).

- [225] B. Treutlein, A. Muschielok, J. Andrecka, A. Jawhari, C. Buchen, D. Kostrewa, F. Hög, P. Cramer, and J. Michaelis, *Dynamic Architecture of a Minimal RNA Polymerase II Open Promoter Complex.*, Mol Cell **46**, 136 (2012).
- [226] K.-J. Armache, S. Mitterweger, A. Meinhart, and P. Cramer, *Structures of complete RNA polymerase II and its subcomplex, Rpb4/7.*, J Biol Chem **280**, 7131 (2005).
- [227] D. Wang, D. A. Bushnell, K. D. Westover, C. D. Kaplan, and R. D. Kornberg, *Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis.*, Cell **127**, 941 (2006).
- [228] H.-T. Chen and S. Hahn, *Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC.*, Cell **119**, 169 (2004).
- [229] H.-T. Chen, L. Warfield, and S. Hahn, *The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex.*, Nat Struct Mol Biol **14**, 696 (2007).
- [230] T. K. Kim, R. H. Ebright, and D. Reinberg, *Mechanism of ATP-dependent promoter melting by transcription factor IIH.*, Science **288**, 1418 (2000).
- [231] H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson, *Structure of a B-DNA dodecamer: conformation and dynamics.*, Proc Natl Acad Sci U S A **78**, 2179 (1981).
- [232] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, *Features and development of Coot.*, Acta Crystallogr D Biol Crystallogr **66**, 486 (2010).
- [233] C. Giardina and J. T. Lis, *DNA melting on yeast RNA polymerase II promoters.*, Science **261**, 759 (1993).
- [234] W. Wang, M. Carey, and J. D. Gralla, *Polymerase II promoter activation: closed complex formation and ATP-driven start site opening.*, Science **255**, 450 (1992).
- [235] G. Miller and S. Hahn, *A DNA-tethered cleavage probe reveals the path for promoter DNA in the yeast preinitiation complex.*, Nat Struct Mol Biol **13**, 603 (2006).
- [236] S. Buratowski, S. Hahn, L. Guarente, and P. A. Sharp, *Five intermediate complexes in transcription initiation by RNA polymerase II.*, Cell **56**, 549 (1989).
- [237] F. W. Martinez-Rucobo, S. Sainsbury, A. C. M. Cheung, and P. Cramer, *Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity.*, EMBO J **30**, 1302 (2011).
- [238] T. R. Strick, J. F. Allemand, D. Bensimon, and V. Croquette, *Behavior of super-coiled DNA.*, Biophys J **74**, 2016 (1998).

Bibliography

- [239] K. Kuznedelov, N. Korzheva, A. Mustaev, and K. Severinov, *Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length.*, EMBO J **21**, 1369 (2002).
- [240] S. Naji, M. G. Bertero, P. Spitalny, P. Cramer, and M. Thomm, *Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement.*, Nucleic Acids Res **36**, 676 (2008).
- [241] B. J. Klein, D. Bose, K. J. Baker, Z. M. Yusoff, X. Zhang, and K. S. Murakami, *RNA polymerase and transcription elongation factor Spt4/5 complex structure.*, Proc Natl Acad Sci U S A **108**, 546 (2011).
- [242] P. Cabart, P. abart, A. Ujvari, M. Pal, and D. S. Luse, *Transcription factor TFIIF is not required for initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIIB in early elongation complexes.*, Proc Natl Acad Sci U S A **108**, 15786 (2011).
- [243] P. F. Kosa, G. Ghosh, B. S. DeDecker, and P. B. Sigler, *The 2.1-Å crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box.*, Proc Natl Acad Sci U S A **94**, 6042 (1997).
- [244] F. T. Tsai and P. B. Sigler, *Structural basis of preinitiation complex assembly on human pol II promoters.*, EMBO J **19**, 25 (2000).
- [245] F. Hayashi, R. Ishima, D. Liu, K. I. Tong, S. Kim, D. Reinberg, S. Bagby, and M. Ikura, *Human general transcription factor TFIIIB: conformational variability and interaction with VP16 activation domain.*, Biochemistry **37**, 7941 (1998).
- [246] S. T. Smale and J. T. Kadonaga, *The RNA polymerase II core promoter.*, Annu Rev Biochem **72**, 449 (2003).
- [247] E. Lehmann, *Structure-function analysis of RNA polymerase II: use of RNA as a template and site-directed mutagenesis in S. cerevisiae.*, PhD thesis, Ludwig-Maximilians-Universität München, Fakultät Chemie und Pharmazie, 2011.
- [248] J. Eichner, H.-T. Chen, L. Warfield, and S. Hahn, *Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex.*, EMBO J **29**, 706 (2010).
- [249] M. Seizl, L. Larivière, T. Pfaffeneder, L. Wenzek, and P. Cramer, *Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization.*, Nucleic Acids Res **39**, 6291 (2011).

- [250] G. Vámosi, C. Gohlke, and R. M. Clegg, *Fluorescence characteristics of 5-carboxytetramethylrhodamine linked covalently to the 5' end of oligonucleotides: multiple conformers of single-stranded and double-stranded dye-DNA complexes.*, *Biophys J* **71**, 972 (1996).
- [251] K. Muller, H. Amman, D. Doran, P. Gerber, K. Gubernator, and G. Schrepfer, *MOLOC: A molecular modeling program*, *Bull. Soc. Chim. Belg.* **97**, 655–667 (1988).
- [252] G. Längst and P. B. Becker, *ISWI induces nucleosome sliding on nicked DNA.*, *Mol Cell* **8**, 1085 (2001).
- [253] Y. Lorch, B. Davis, and R. D. Kornberg, *Chromatin remodeling by DNA bending, not twisting.*, *Proc Natl Acad Sci U S A* **102**, 1329 (2005).
- [254] A. Saha, J. Wittmeyer, and B. R. Cairns, *Chromatin remodeling by RSC involves ATP-dependent DNA translocation.*, *Genes Dev* **16**, 2120 (2002).
- [255] A. Saha, J. Wittmeyer, and B. R. Cairns, *Chromatin remodelling: the industrial revolution of DNA around histones.*, *Nat Rev Mol Cell Biol* **7**, 437 (2006).
- [256] D. Vasudevan, E. Y. D. Chua, and C. A. Davey, *Crystal structures of nucleosome core particles containing the '601' strong positioning sequence.*, *J Mol Biol* **403**, 1 (2010).
- [257] M. Holzner, *Nucleosome remodeling by Chd1.*, Master's thesis, Ludwig-Maximilians Universität München, 2010.
- [258] K. Rippe, A. Schrader, P. Riede, R. Strohner, E. Lehmann, and G. Längst, *DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes.*, *Proc Natl Acad Sci U S A* **104**, 15635 (2007).
- [259] J. G. Yang, T. S. Madrid, E. Sevastopoulos, and G. J. Narlikar, *The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing.*, *Nat Struct Mol Biol* **13**, 1078 (2006).
- [260] N. H. Thomä, B. K. Czyzewski, A. A. Alexeev, A. V. Mazin, S. C. Kowalczykowski, and N. P. Pavletich, *Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54.*, *Nat Struct Mol Biol* **12**, 350 (2005).
- [261] M. G. Pray-Grant, J. A. Daniel, D. Schieltz, J. R. Yates, 3rd, and P. A. Grant, *Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation.*, *Nature* **433**, 434 (2005).
- [262] T. Sengoku, O. Nureki, A. Nakamura, S. Kobayashi, and S. Yokoyama, *Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa.*, *Cell* **125**, 287 (2006).

Bibliography

- [263] S. Hohng, C. Joo, and T. Ha, *Single-molecule three-color FRET.*, Biophys J **87**, 1328 (2004).
- [264] B. Huang, M. Bates, and X. Zhuang, *Super-resolution fluorescence microscopy.*, Annu Rev Biochem **78**, 993 (2009).
- [265] J.-M. Rouillard, W. Lee, G. Truan, X. Gao, X. Zhou, and E. Gulari, *Gene2Oligo: oligonucleotide design for in vitro gene synthesis.*, Nucleic Acids Res **32**, W176 (2004).
- [266] K. Luger, T. J. Rechsteiner, and T. J. Richmond, *Expression and purification of recombinant histones and nucleosome reconstitution.*, Methods Mol Biol **119**, 1 (1999).
- [267] A. Gansen, K. Tóth, N. Schwarz, and J. Langowski, *Structural variability of nucleosomes detected by single-pair Förster resonance energy transfer: histone acetylation, sequence variation, and salt effects.*, J Phys Chem B **113**, 2604 (2009).
- [268] C. Claudet, D. Angelov, P. Bouvet, S. Dimitrov, and J. Bednar, *Histone octamer instability under single molecule experiment conditions.*, J Biol Chem **280**, 19958 (2005).
- [269] D. Rhodes and R. A. Laskey, *Assembly of nucleosomes and chromatin in vitro.*, Methods Enzymol **170**, 575 (1989).
- [270] S. Awad, D. Ryan, P. Prochasson, T. Owen-Hughes, and A. H. Hassan, *The Snf2 homolog Fun30 acts as a homodimeric ATP-dependent chromatin-remodeling enzyme.*, J Biol Chem **285**, 9477 (2010).
- [271] G. Meersseman, S. Pennings, and E. M. Bradbury, *Mobile nucleosomes—a general behavior.*, EMBO J **11**, 2951 (1992).
- [272] R. E. Benesch and R. Benesch, *Enzymatic removal of oxygen for polarography and related methods.*, Science **118**, 447 (1953).
- [273] I. Rasnik, S. A. McKinney, and T. Ha, *Nonblinking and long-lasting single-molecule fluorescence imaging.*, Nat Methods **3**, 891 (2006).
- [274] T. Cordes, J. Vogelsang, and P. Tinnefeld, *On the mechanism of Trolox as anti-linking and antibleaching reagent.*, J Am Chem Soc **131**, 5018 (2009).
- [275] J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, Jr, and C. H. Wiggins, *Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data.*, Biophys J **97**, 3196 (2009).

-
- [276] A. J. Andrews, X. Chen, A. Zevin, L. A. Stargell, and K. Luger, *The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions.*, Mol Cell **37**, 834 (2010).
- [277] M. J. Rust, M. Bates, and X. Zhuang, *Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).*, Nat Methods **3**, 793 (2006).
- [278] S. W. Hell, *Far-field optical nanoscopy.*, Science **316**, 1153 (2007).
- [279] X. Darzacq, J. Yao, D. R. Larson, S. Z. Causse, L. Bosanac, V. de Turris, V. M. Ruda, T. Lionnet, D. Zenklusen, B. Guglielmi, R. Tjian, and R. H. Singer, *Imaging transcription in living cells.*, Annu Rev Biophys **38**, 173 (2009).
- [280] G.-W. Li and X. S. Xie, *Central dogma at the single-molecule level in living cells.*, Nature **475**, 308 (2011).
- [281] F. Cella Zanacchi, Z. Lavagnino, M. Perrone Donnorso, A. Del Bue, L. Furia, M. Faretta, and A. Diaspro, *Live-cell 3D super-resolution imaging in thick biological samples.*, Nat Methods **8**, 1047 (2011).
- [282] J. P. Siebrasse, T. Kaminski, and U. Kubitscheck, *Nuclear export of single native mRNA molecules observed by light sheet fluorescence microscopy.*, Proc Natl Acad Sci U S A **109**, 9426 (2012).
- [283] T. A. Planchon, L. Gao, D. E. Milkie, M. W. Davidson, J. A. Galbraith, C. G. Galbraith, and E. Betzig, *Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination.*, Nat Methods **8**, 417 (2011).

Acknowledgements

I would like to express my deep gratitude to all those that contributed to the success of this work. This thesis would not have been possible without the support of numerous people.

Foremost, I want to thank my supervisor Prof. Jens Michaelis for giving me the opportunity to pursue such exciting basic biological questions and for continuously supporting me throughout the past years. His way of thinking about science and tackling problems strongly influenced my scientific development during my PhD. I want to thank him for his positive attitude and the stimulating and encouraging discussions, which were crucial for keeping me motivated and turning my efforts into success. I found it special that Jens was not only interested in the scientific success, but also cared about the personal life of his PhD students and allowed me to work at times from across the ocean. Jens, thank you for this trust and personal support!

I particularly want to thank Prof. Patrick Cramer for the exciting and fruitful collaboration on Pol II transcription initiation. I am extremely grateful for all the stimulating discussions and for everything I learned from him about Pol II structure and function. Furthermore, I am very grateful for his help with our manuscript - I learned a lot about how to write a good paper from him. Moreover, I want to thank him for being "Zweitgutachter" of this thesis.

Several members of the Cramer lab owe special thanks. I want to thank Claudia Buchen for all Pol II-F preparations and her supportive attitude, Dirk Kostrewa for modeling the Pol II open complex, and Rike Hoeg for the nice discussions, in vitro transcription assays and the enjoyable time at CSH! I want to thank Anass Jawhari, my contact person in the Cramer lab in the very beginning, for his enthusiasm about the PIC project, helpful discussions and the fun time together with Joanna (Asia) Andrecka as the "PIC team". Moreover, Alan Cheung for help with modeling in coot, Kerstin Kinkelin for preparing recombinant TFIIF and Sarah Sainsbury for nice discussions and Pol II.

Essential for this thesis was a collaboration with the group of Prof. Tom Owen-Hughes on Chd1 remodeling. I want to thank Tom Owen-Hughes for this exciting collaboration and for stimulating discussions. Moreover, I am deeply thankful to his former Postdoc Daniel Ryan for his dedication to this project, for all the Chd1 preparations and for all the stimulating discussions and his input into the Chd1 project.

Along this line, I want to thank Prof. Gernot Längst for our collaboration in the beginning of this thesis. He introduced me into the world of nucleosomes and chromatin remodelers. My special gratitude goes to his former PhD student Josef Exler, who taught me all the ins and outs of nucleosome assembly and remodeling assays. I very much enjoyed our collaboration and it is a pity, that ACF never behaved so well in the smFRET experiments.

After having considered all my collaborations, I want to return to our own floor in the physical chemistry building. I want to thank Professor Christoph Bräuchle for always being motivating

Bibliography

and inspiring to his students, and Prof. Don Lamb for pleasant and stimulating discussions. It was a very special experience and great pleasure for me to be part of the unique work atmosphere created by the Bräuchle, Lamb and Michaelis group, an atmosphere both extremely pleasant and very productive.

With this I want to thank all current and former members of the three labs for contributing to such a unique and synergistic work environment, for always being open for discussions and always finding time to help.

I am sincerely thankful to Asia (Joanna Andrecka), who introduced me into biochemistry, smFRET and NPS analysis of Pol II complexes in the beginning of my doctoral work. Dzieki Asiu, for all this enjoyable time together in the lab, in the mountains, on travels, all the coffees, salads, glasses of wine... and for infecting me with your enthusiasm for traveling.

I want to thank Adam Muschielok for his development of the sophisticated NPS analysis software, which was essential for this thesis. I am deeply thankful for his efforts to constantly modify and improve the program such that it met the needs of my projects. Moreover, I want to thank him for his never-ending curiosity that initiated numerous enjoyable and enriching discussions. Dzieki bardzo Adam!

I am extremely thankful to Robert Lewis, who built and continuously developed further the TIRF microscope that I used to acquire all my smFRET data and who developed the software to analyze smFRET movies. Moreover I want to thank Robert for his helpfulness - Robert always took the time to answer questions, even after having left the group to his new job.

Another software I used in this thesis for the examination of dynamic smFRET time traces and for HMM analysis of such traces was written by Gregor Heiss with the help of Martin Sikor. I am very thankful to Gregor and Martin for the development and constant improvement of this software and for all the help with Hidden Markov Modeling and Matlab in general.

Along this line, I want to thank Wolfgang Kügel and Jochen Reichel for their help with technical questions and computer problems.


Many people helped me by expressing and purifying proteins that I used in the smFRET measurements. I want to thank Katarzyna (Kasia) Krzemień and Monika Holzner for the preparation of recombinant histones and Jarka Obel for preparing TFIIB and TBP.

My special gratitude goes to Ania Jasiak for her efforts in developing a strategy for labeling TFIIB and for teaching my how to express and purify proteins on my own. Moreover, I want to thank Ania for all the support and deep conversations during coffee breaks and the salad lunches! Asia i Ania (i Grofa), dzieki bardzo dla ciepły dom w Allgäuer Strasse!

I want to thank Monika Holzner for joining in and continuing my research on Chd1 remodeling. Moni, thank you for all the DNA preparations, nucleosome assemblies and the help with smFRET measurements. Moreover, I want to thank Monika Holzner and Julia Nagy for the great, enjoyable company in the lab and beyond, for the countless chamber days and for always cheering me up. I also want to thank my Bachelor- and Master-students that worked with me over the past years: Johanna Ludwigsen, Monika Holzner, Anne-Kathrin Witt, Marta Wesolowska.

I am deeply thankful to my parents and all my siblings with their families for their support and for being such great family! In particular, I want to thank Philipp and Elke for the enjoyable moments together in München and for all the support and advice, Gudi for the pleasant time living together and for her lighthearted nature, and Hilde for always being there for me.

Finally, I want to thank Gray for his love and his support through all means of communication, for his endurance with me, his "quirliness", which always cheers me up and for his miasmal enthusiasm for nature and biology! :-)



This thesis focuses on understanding the molecular mechanisms of two fundamental processes of eukaryotic gene expression: transcription initiation and nucleosome remodeling. Since both processes are characterized by large conformational changes and a high flexibility of nucleoprotein complexes, standard high-resolution structural methods are hindered and instead direct visualization in real time is required as provided by single molecule techniques.

In the first part of this work, I used single molecule fluorescence resonance energy transfer (smFRET) experiments, Nano-Positioning System (NPS) analysis and x-ray crystallographic information to determine the three-dimensional architecture of a minimal Pol II open promoter complex (OC) consisting of promoter DNA, TBP, Pol II and general transcription factors TFIIB and TFIIF. In the OC, TATA-DNA and TBP reside above the Pol II cleft between clamp and protrusion domains. The TFIIB core domain is displaced from the Pol II wall, where it is located in the closed promoter complex. Furthermore, I directly observed the downstream DNA to be dynamically loaded into and unloaded from the Pol II cleft at a timescale of seconds. These results uncover large overall structural changes during the initiation-elongation transition.

In the second part, I applied the same experimental approach to determine the location of the three constituting domains of chromatin remodeler Chd1 (Chromodomain-helicase-DNA-binding protein 1) in a Chd1-nucleosome complex that exhibited intrinsic dynamics. The NPS results allowed me to construct a model of the Chd1-nucleosome complex, in which the DNA-binding domain is associated with extranucleosomal DNA at the nucleosome entry site, the tandem chromodomains are located below the entry site close to histone H4 tail and the ATPase motor binds nucleosomal DNA between dyad and superhelical location +1. Furthermore, I used smFRET to follow in real-time the structural dynamics of nucleosomal DNA during Chd1 catalyzed repositioning. FRET time trajectories revealed gradual and bidirectional translocation of nucleosomal DNA by Chd1 and the data allowed me to propose a model for the remodeling mechanism of Chd1, which involves formation and propagation of a DNA loop.

In summary, the results of this work present a substantial progress in our understanding of the molecular mechanisms of transcription initiation and nucleosome repositioning and they demonstrate that smFRET combined with NPS analysis is a very powerful tool to study the structure as well as the dynamics of large, flexible nucleic acid-protein complexes, not amenable to structural and biochemical ensemble techniques.

